The Mechanism of RecA Mediated DNA Patterning Interrogated by AFM

Andrew James Lee

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Centre for Doctoral Training in Molecular Scale Engineering
School of Electronic and Electrical Engineering

January 2016

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

CHAPTER 4: is derived from a book chapter published by the candidate (primary author) and prof. C. W$\ddot{a}$lti (co-author) - giving a detailed review of AFM mechanics and application in bionanomaterials research.

Publication: Studying biologically template materials with atomic force microscopy, Lee, A.J. & W$\ddot{a}$lti, C. in Nanomaterials: A Guide to Fabrication and Applications (Krishnamoorthy, S.) CRC Press, 2015. ISBN: 978-1-4665-9125-7

CHAPTER 7: contains materials published by the candidate and co-authors (including Prof J. K. Hobbs and Prof. C W$\ddot{a}$lti both of whom are the candidates supervisors). The work was conducted by the candidate, with contribution from Dr. M. Szymonik to the coding of the python script, which was used to compare vector representations of DNA molecules. This contribution is acknowledged in the thesis text.

Publication: Lee, A. J. Szymonik, M. Hobbs, J. K. and W$\ddot{a}$lti, C. Tuning the Translational Freedom of DNA for High Speed AFM. Nano Research, 8(6):1811-1821, (2015).

Abstract

Over recent years, advancements in bottom - up construction technologies are enabling the creation of heterogeneous and functional materials. These approaches offer the potential to surpass the physical limitations in traditional top - down micromachining. Of these, bionanotechnological approaches that harness the inherent molecular recognition and self-assembling properties of biological molecules - such as Deoxyribose nucleic acid (DNA) - are arguably the most promising.

Over recent years, the field of DNA nanotechnology has advanced rapidly, enabling the creation of arbitrary structures in two and three dimensions. These substrates act as adapters enabling the arrangement of functional components at the nano-scale to be interfaced with the macro-scale world.

One approach to spatially address DNA nano-architectures is to harness the sequence specific homologous recombination mechanism of the *E.coli* protein Recombinase A (RecA). This protein mediates the alignment of a supplied single stranded DNA (ssDNA) with a subject double stranded DNA (dsDNA) where homology is shared, making this method inherently programmable. Despite several successful demonstrations of the artificial application of RecA, the underlying mechanism which orchestrates this interaction remains widely debated. The lack of clear understanding surrounding this critical biological mechanism stems from the in-direct approaches taken to interrogate it, to date.

In response to this, the work presented in this thesis, attempts to answer the open biological questions surrounding RecA. Here, recent advances in high speed atomic force microscopy (HS-AFM) and high resolution Atomic force microscopy (AFM) - using rapid-force-curve imaging - are applied to directly interrogate the homology searching mechanism of RecA. When taken together, these structural and functional insights will inform the future development of RecA mediated patterning approaches within complex DNA topologies.

This page is intentionally left blank.

This work is dedicated to those of us who thought
we would never achieve anything.

This page is intentionally left blank.

” I learned very early the difference between knowing the name of something and knowing something. ”

– *Richard Feynman*

This page is intentionally left blank.

Acknowledgements

I would like to express my sincerest gratitude to Prof. Christoph Wälti and Prof. Jamie Hobbs (University of Sheffield) for their excellent guidance, undying patience and immeasurable expertise, without which this project would not have been possible. I would further like to extend my warmest appreciation to Dr Rajan Sharma and Dr Mike Szymonik for their invaluable advice, support and continued input to this work.

Additionally, I would like to thank the many collaborators who have contributed their time and expertise to this project. In particular; Prof. Masayuki Endo (Kyoto University) for his help in the development of a larger DNA origami frame and other on going work; Dr Bede Pittenger, Dr Andrea Slade and Dr Thomas Muller (Bruker Nanosurfaces division, Santa Barbara, CA, USA) and Dr Hartmut Stadler (Bruker Nanosurfaces division, Karlsruhe, Germany) for their thoughtful discussion and support regarding optimising Peakforce tapping and HS-AFM while on location in Santa Barbara and Karlsruhe, respectively.

Furthermore, I feel I must highlight the motivational and stimulating atmospheres of both the Bioelectronics group and the CDT for molecular scale engineering, whose personal and professional impacts are too numerous to list. In particular, I wish to thank my close friends; Will Morton, Sybilla Corbett, Oliver Chammas, Mark Sellick and Scott Bird with whom I have shared the best and worst this PhD process had to offer.

Finally, I'd like to thank my partner; Marion for being my rock throughout this PhD, providing the emotional support and encouragement I have needed to complete this work. Along with my mother, father and brother who are my continued source of support and inspiration. I'd also like to acknowledge all my friends who have kept the light in my life throughout the darkest days of this PhD.

This page is intentionally left blank.

Contents

Abbreviations

T_m Melting temperature.

AFM Atomic force microscopy.

bp Base pair.

DNA Deoxyribose nucleic acid.

dNTPs Deoxynucleotide Triphosphates.

dsDNA double stranded DNA.

HS-AFM high speed atomic force microscopy.

Kbp Kilobase pair.

MW Molecular weight.

nm Nanometre.

NPF Nucleoprotein filament.

PCR Polymerase chain reaction.

RecA Recombinase A.

RNA Ribose nucleic acid.

ssDNA single stranded DNA.

TEM Transmission electron microscopy.

This page is intentionally left blank.

List of figures

This page is intentionally left blank.

Chapter 1

Introduction

In recent years, materials science has been re-invigorated by advancements made in the field of nanotechnology. However, to date, such developments typically allow for the uniform formation of homogeneous materials, offering enhanced characteristics such as strength. In contrast, in order to form functional nano-scale architectures it is necessary to order materials heterogeneously in a specifically coordinated manner.

To this end, bottom-up construction technologies that harness the inherent molecular-recognition and self-assembling properties of biological systems have the potential to surpass the most advanced top-down microfabrication techniques, which are reaching the physical limits of the current technology.

It can be argued that DNA nanotechnology (see section 3.1) presents the most promising route to realising the rational design of devices from biological molecules. Here, DNA is considered for its structural properties and the fidelity of its base-pairing interactions, enabling the construction of branched junctions and higher order structures in two and three dimensions.[1] This field has seen rapid growth from individual junctions produced from several ssDNA molecules, to highly efficient scaffolded approaches that weave viral DNA to form arbitrary structures - referred to as DNA origami (see section 3.1.3).[2]

Moreover, recent advances in the directed surface organisation of DNA structures hails in an era where biologically derived systems formed in solution may be integrated directly with micro-fabricated solid state structures.[3] Hence, DNA origami may be utilised as a molecular breadboard upon which components are arranged to form functional subunits. These subunits can be further arranged across larger length scales to form integrated circuits and functional devices. Therefore, DNA nanotechnology maybe considered an adapter, able to interface nanoscale arrangements with

the macroscale world.

In order to orchestrate functions, the heterogeneous arrangement of components to infer spatial asymmetry, polarity or directionality is required. In some instances, these arrangements can be pre-designed into the DNA structure through the inclusion of uniquely functionalised ssDNA molecules or included indirectly through the introduction of different DNA binding proteins – which themselves are chemically functionalised.

These methods are, however, relatively inflexible where the number of uniquely addressable protein binding sites is typically very low within the available sequence space. Further, any change in desired functionality would require a re-design of the entire DNA assembly, making these approaches rather costly.

Alternatively, an approach being developed in the Bioelectronics group at the University of Leeds, is to harness the homologous recombination mechanism of the *E.coli* protein, RecA - commonly referred to as RecA mediated patterning.[4,5,6] The RecA protein contains two DNA binding sites, thus able to accommodate ssDNA and dsDNA simultaneously mediating a triple stranded complex.

In this scheme, RecA monomers polymerise on ssDNA in the presence of ATP and $Mg^{2+}$to form a nucleoprotein filament (NPF) (figure 1.1). Subsequently, this NPF searches for a region of sequence homology between the encapsulated ssDNA and a target dsDNA molecule. Where homology is located RecA mediates the formation of a triple-stranded DNA complex and undertakes a strand exchange - recombining the DNA molecules to form a new heteroduplex dsDNA - following which the complex disassembles.[7]

For the purposes of bionanotechnology, this process can be arrested at the triple stranded DNA stage where NPFs are formed with ATPγS - a non-hydrolysable analogue of ATP. (figure 1.1).[4,5]

As such, locations on a DNA substrate may be targeted arbitrarily and specifically, without the need to re-design the underlying template. For example, a single DNA origami tile design may be addressed in many different ways to form a multitude of different devices (figure 1.1), hence making this approach highly economical.

To date, bionanotechnological applications of RecA have been successfully demonstrated, such as the formation of conductive nanowires on linear and branched DNA templates.[8] Furthermore, RecA has been used to form field effect transistors (FET) through the organised placement of carbon nanotubes[9] and as a molecular lithography mask to create insulating gaps in conductive

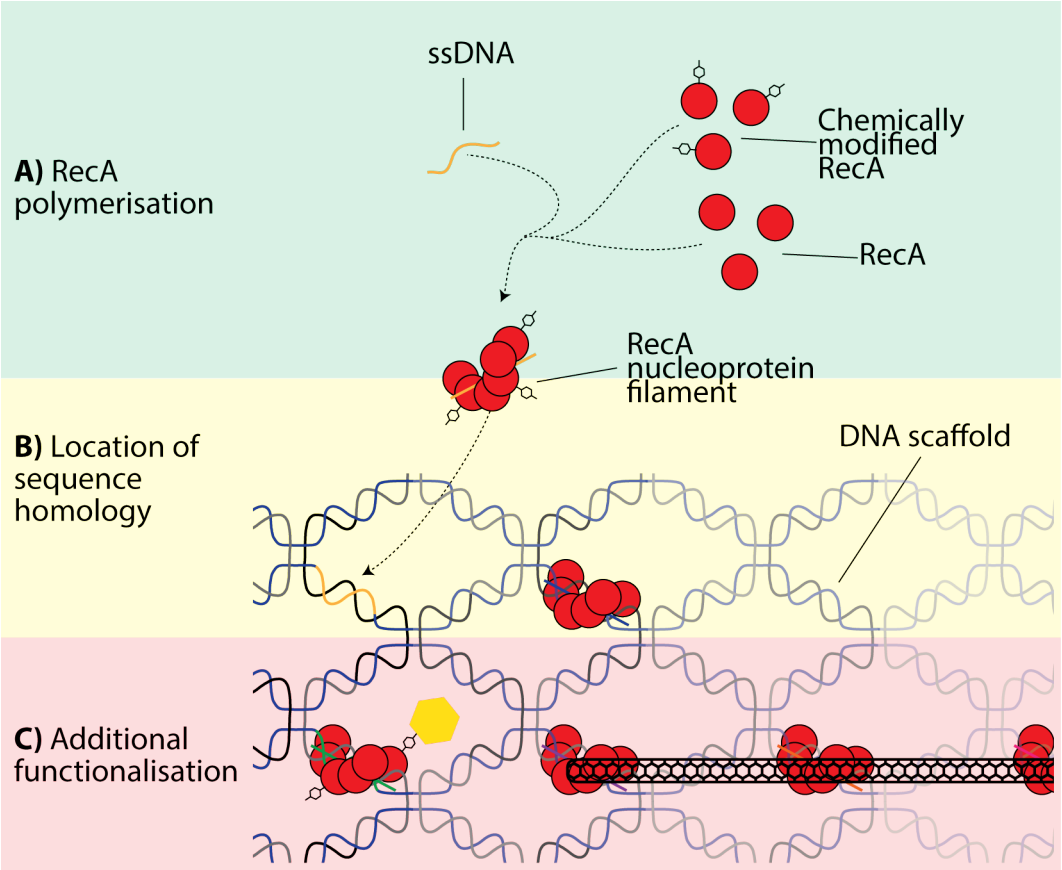

Figure 1.1: ***A schematic overview of the envisaged RecA mediated hierarchical construction on DNA scaffolds.*** *DNA structures are spatially addressed using the homologous recombination ability of RecA.* ***A)*** *Firstly, RecA or chemically modified RecA is polymerised onto ssDNA (orange) to form a nucleoprotein filament (NPF).* ***B)*** *NPFs seek out regions of sequence homology (shown in orange) between the encapsulated ssDNA and the complex dsDNA template leading to site specific patterning.* ***C)*** *RecA is used to add additional functional components through chemical modification, such as a conductive circuits (e.g. carbon nanotubes) or photonic elements (e.g. gold nanoparticles).*

nanowires.[10] Further work in the Bioelectronics group has demonstrated that this approach can be successfully multiplexed and scaled down to allow feature sizes of less than 10 nm.[4,5,6]

However, despite the success of these studies, severe gaps remain in the current understanding of the mechanism by which RecA mediates homologous recombination. In particular, how NPFs are able to conduct a homology search remains widely debated. Current opinion is divided as to whether NPFs are able to conduct a one dimensional facilitated diffusion along dsDNA, possibly favouring random sampling by three dimensional diffusion. Further proponents have provided evidence for the involvement of intersegmental transfer.[11] It is likely that the true mechanism is a combination of all these interactions.

This lack of clear understanding can be largely attested to the methods of study employed to date, which typically rely on in-direct reporting [12,13,14,11,15,16,17] or have attempted to infer functional intermediates from ensemble structural derivation methods – such as crystallography. [18,19,20,21,22] Such limited insight has implications for the bionanotechnological application of RecA where it would be employed in environments that far exceed its physiological role. For example, how would the presence of complex DNA topology such as the woven structure of DNA origami impact the homology searching ability of RecA NPFs?

1.1 Project Aim

In light of this, the work presented in this thesis attempts to directly investigate the RecA mediated homology searching mechanism with a combination of HS-AFM and high spatial resolution AFM.

Here, the primary mode of interaction between NPFs and dsDNA is directly observed in real time with HS-AFM. More specifically, this project aims to identify and quantify the suggested existence of a facilitated diffusion mechanism directly for the first time.

In order to achieve this, the work presented here aims to develop a unique sample preparation method to enable a transient binding of dsDNA sufficient for observation with the HS-AFM whilst accommodating NPF interactions.

This must also be accompanied by a statistical analysis of homology searching intermediates, in order to reconcile any real time observations and take account of sample preparation or AFM induced artefacts.

In addition, the local manipulation of DNA structure as a result of NPF interaction is investigated to characterise any induced strain which may be limiting in constrained DNA topologies. These high resolution observations – made using rapid-force-curve imaging (see section 4.4.2) – aim to increase the spatial resolution on the isolated biological complexes – NPF and DNA structures – to examine any structural variation therein which may support the observed real-time dynamic interactions.

1.2 Chapters Outlines

With respect to the aims of this project, this section outlines the content of the following chapters in this thesis.

Chapter 2

Given the project outline detailed above, this chapter explores the biological molecules DNA and proteins, that form the core of these bionanotechnological approaches. These critical biological molecules are introduced, describing their biological role, their basic subunit chemistry and the final physiological geometries adopted. This sets the scene for a discussion of how they can be manipulated for artificial applications in the following chapter.

Chapter 3

This chapter considers the extant literature on DNA nanotechnology and RecA in justification of the RecA mediated patterning approach described above. In the first instance, the inception and development of DNA nanotechnology is considered, with reference to the specific characteristics of DNA that make it uniquely suited for this application. Further, the development of DNA origami approaches is detailed and its growth into three dimensions is examined. Finally the emerging applications for this technology after three decades of development is considered.

The second half of this chapter examines RecA in relation to its application to spatially address DNA nano architectures. The structure and biological roles of the RecA monomer and its polymer - the NPF - are examined. Furthermore, the current published literature regarding the non-native application of RecA is examined within the context of the RecA mediated patterning mechanism described here.

Finally, methods used to interrogate the RecA homology searching mechanism to date are explored and used to justify the use of direct imaging approaches with the AFM in this thesis.

Chapter 4

As the AFM is the core instrument used throughout this project a thorough examination of the underpinning physics of its operation is given in this chapter. A brief history of the development of AFM is first explored followed by a detailed discussion of the tip sample interactions and cantilever dynamics.

The proceeding sections go on to define the spatial resolutions, their relation to the force sensitivity and relevant developments in rapid-force-curve imaging. Moreover, the temporal resolutions of the HS-AFM are examined and related instrument developments are described.

Chapter 5

The reader is provided with a detailed overview of the materials and methods utilised throughout the proceeding chapters, in particular; Polymerase chain reaction (PCR), gel electrophoresis, RecA mediated patterning and associated restriction enzyme assay.

Chapter 6

In the first experimental chapter, work to statistically assess the intermediates of a typical NPF interaction with DNA is detailed. Here, a 60 minute reaction window is temporally dissected at $37°C$ and $22°C$. Moreover, the effect of heterology and NPF size are explored. Finally, how the reaction landscape is altered due to the presence of a solid-liquid interface is detailed in relation to the patterning NPFs in constrained and artificial environments.

Chapter 7

This chapter proceeds to describe the development of a sample preparation for orchestrating a transient surface absorption of dsDNA molecules. This is described alongside a methodology for quantifying the translational freedom of DNA molecules, which when taken together enable the mobility of DNA molecules adhered to a mica surface to be tuned appropriately to accommodate nucleoprotein interactions.

Further, the utility of this approach is demonstrated using the restriction enzyme, EcoRI as a proof of principle. Subsequently, initial observations of RecA polymerisation and NPF homology searching on dsDNA are described.

Chapter 8

In light of limitations imposed by the previous experimental design, the work in this chapter adopts DNA origami as a reference structure from which positional and orientational information regarding the observed interactions can be quantified.

The formation of these DNA origami structures is described followed by the significant observations made therein. The latter part of this chapter describes the collaborative efforts between the author and the Sugiymama group (Kyoto University, Kyoto, Japan) to further enlarge the DNA origami reference structure to enhance the experimental setup.

Chapter 9

The final results chapter explores the application of rapid-force-curve imaging to surpass the published spatial resolutions on DNA, DNA origami and RecA NPFs. This work provides important insights regarding unique local structural variations, which ensemble methods are insensitive to. Furthermore, attempts to apply this imaging methodology to directly observe the structural manipulation of dsDNA by NPFs are detailed.

Chapter 10

This final chapter, summarises the results achieved in this thesis laying them in context with the wider field. The significant implications of the observations made here are explored and future directions of this work are described.

This page is intentionally left blank.

Chapter 2

Introduction to DNA and Proteins

Biological molecules are appealing candidates for the development of structural and functional nanotechnology. First and foremost, nature already successfully forms and operates a multitude of intricate and complex machines at these length scales.[23] Inherent to this is the process of self-assembly.

This governing phenomena is the implicit result of the constituent molecules carrying all the required information to form the final structure. As a consequence, the majority of molecules can spontaneously form and reform their designated geometry without external direction or input.

It is therefore easy to see how this is an appealing concept for the formation of synthetic nanoscale devices. Of the abundant biological molecules that exist, two immediately present themselves as ideal candidates for the production of synthetic objects; proteins and nucleic acids - more specifically DNA.[24] Naturally, the former typically forms all the functional machines within a cell and the latter stores and transmits information through the specificity of its internal interactions.

In this chapter the form and function of these critical biological components - DNA and protein - is described, setting the context for their exploitation in bionanotechnological applications, such as the work presented in this thesis.

2.1 Deoxyribonucleic Acid (DNA)

All life depends on the ability to store, translate and transmit genetic information. This can be described by the central Dogma of Biology which discusses the processing of information from its storage in the genome - as DNA - through transcription in to RNA and finally its translation into

functional proteins.

Originally, it was difficult to conceive how DNA might carry the genomic information due to its apparent chemical simplicity, with only four very similar subunits; called nucleotides. It was only with the discovery that DNA consisted of two strands that its capability to store and replicate information was understood. In this section the chemical and functional properties of nucleic acids are described, starting from the basic subunits; the nucleotide.

2.1.1 Nucleic Acid Chemistry

Nucleotides are constructed from a pentose backbone, a nitrogenous base and up to three phosphate groups (figure 2.1).

The backbone pentose molecule - a five-carbon sugar ring (figure 2.1, blue box) is essentially the same amongst all nucleic acids. A subtle change in this sugar backbone is the defining difference between Ribose nucleic acid (RNA) and DNA; β-D-ribose and β-D-2-deoxyribose, respectively. These two differ due to the loss of a OH group, indicated in red in figure 2.1, blue box.[23]

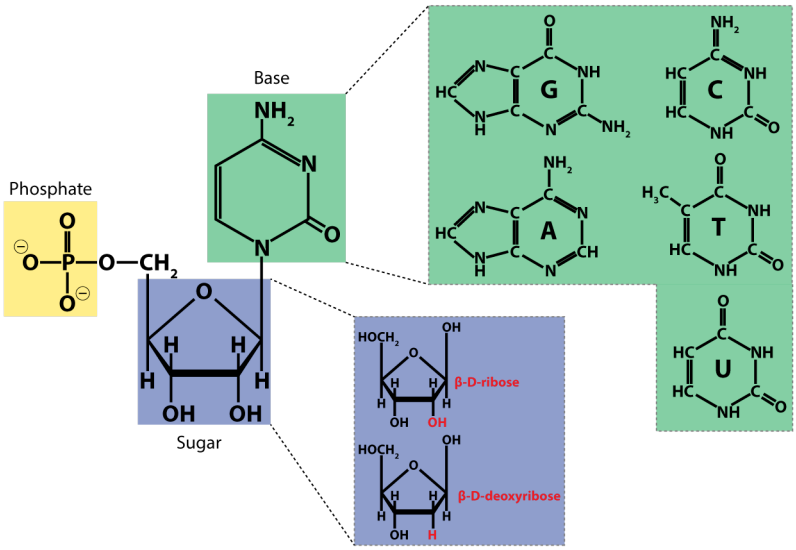

Figure 2.1: ***The chemical structure of nucleic acids.*** *A nucleic acid consists of three major parts, a pentose backbone (blue), a nitrogenous base (green) and a set of phosphate groups (yellow).****(Inset blue box)*** *RNA (top) and DNA (bottom) differ due to the loss of an hydroxyl group (red).* ***(Inset green box)*** *The structure of the five different bases; guanine (G), cytosine (C), adenine (A), thymine (T) and uracil (U) are shown in their respective pairs.*

There are five nitrogenous bases found in nature; cytosine (C), thymine (T), adenine (A), guanine

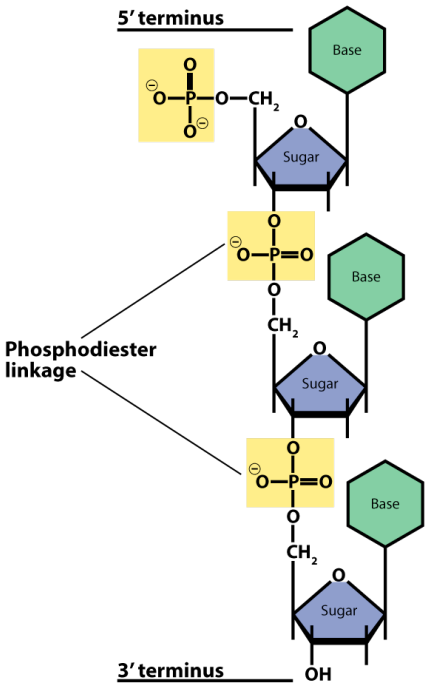

Figure 2.2: ***The chemical structure of the DNA backbone.*** *Nucleic acids link together through the formation of phosphodiester bonds between the phosphate group (yellow) of one molecule and the pentose hydroxyl (blue) group of its neighbour. The resulting oligonucleotide contains an inherent polarity with the terminal phosphate and hydroxyl groups at the 5' and 3', respectively.*

(G) and uracil (U) (figure 2.1, green box). The first 4 bases are utilised in DNA.[25] In comparison, RNA utilises U instead of T. The bases are split into two categories based upon the nature of their carbon rings. Adenine and Guanine are both purines and contain a double ringed structure, in contrast to the singularly ringed pyrimidines (thymine, uracil and cytosine)(figure 2.1).[23]

The final component of a nucleotide is one or more phosphate groups. These are typically bound to the hydroxyl group, indicated in yellow in figure 2.1. A single phosphate group is found in nucleic acids, which is used to form the linkage between subsequent nucleotides in a sequence. In contrast, two and three phosphate groups are found on Adenosine Di (ADP) and Triphosphate (ATP), respectively. These are one of the most abundant energy currencies within Biology, releasing large amounts of energy upon removal of a phosphate group.[23]

Nucleotides are linked together through a phosphodiester bond that forms between the 5' and 3' carbon atoms to form a linear nucleic acid polymer (figure 2.2). Due to the inherent chirality of the constituent nucleotides, the nucleic acid strand exhibits a polarity, with a 3' hydroxyl terminus and a 5' phosphate terminus. Conventionally, DNA sequences are written 5' to 3'. As a result, it is important to note that all the bases are arranged in the same orientation along the length of the polynucleotide chain.[23]

In addition, the sugar-phosphate backbone confers a strong negative charge as a result of the ionisation of the phosphate groups at biological pH, causing a single negative charge to be shared across the phosphodiester bonds between each deoxyribose sugar, making the overall structure

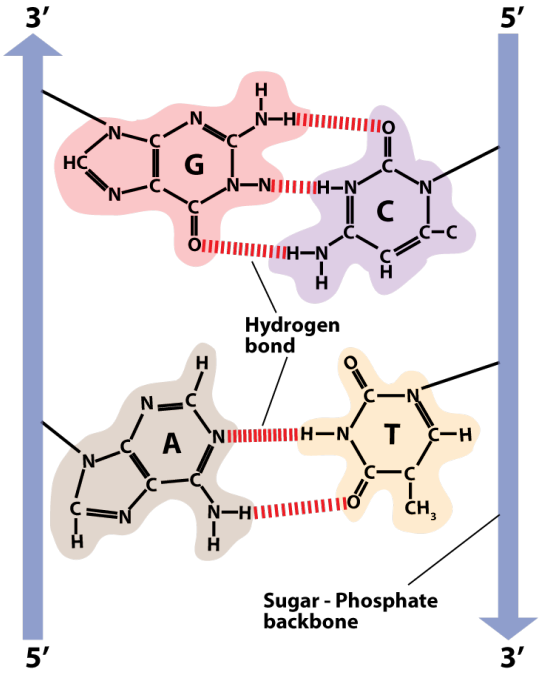

Figure 2.3: ***A schematic diagram depicting the Watson and Crick base pairing regime.*** *G binds to C and A binds to T making 3 and 2 hydrogen bonds (red dashed lines), respectively. The two strands in the resulting duplex DNA run anti-parallel to one and other.*

negative.[26]

2.1.2 The Double Helix

It has been known since 1953, that the native structure of DNA contains two nucleic acid strands.[24] These strands wrap around each other to form the characteristic double helix, as evidenced by the X-ray diffraction studies conducted by Rosalind Franklin in 1950.

In this conformation the bases are found internally and the sugar backbone of both strands faces outwards to the surrounding aqueous environment. This allows the bases to undergo hydrogen (H) bonding with those on the other nucleic acid strand. It is important to note that in order for the bases in the two nucleic acid strands to align and undergo H bonding the strands run in opposing directions, giving rise to an anti-parallel DNA duplex.

In principle every nucleotide is capable of H bonding to every other, giving rise to the Hoogsteen base pairing regime.[27] However, the most energetically favourable complementary base-pairing regime was first described by Watson-Crick; where Adenine (A) bonds to Thymine (T) and Guanine (G) bonds to Cytosine (C), forming two and three hydrogen bonds, respectively (figure 2.3).[24] The sum of these interactions between two ssDNA oligonucleotide dictates the Melting temperature (T_m) of any resultant dsDNA molecule, where the T_m is defined as the point at which 50% of the duplexes are found to have dissociated.[28] As a consequence of this base-pairing regime,

each DNA molecule contains two anti-parallel DNA strands that are exactly complementary to one another. This allows for built in redundancy as one strand maybe used as a template to reproduce the other. It is this predictable binding - and hence base sequence - characteristics that enable the coding of genetic information.

As discussed, the nucleic acid backbone holds a net negative charge in an aqueous environment[26] causing repulsion between neighbouring ssDNA strands – an effect which would prevent the formation of dsDNA. However, under physiological conditions this repulsive force is quenched by the interaction of an excess of cationic species and thus allows the two DNA strands to interact closely enough for base-pairing to occur.[23] It is therefore important to consider that the salt concentration has a large impact on the melting temperature (T_m) of dsDNA, in addition to its base sequence. It is found that a $17°C$ decrease in T_m occurs per 10 fold decrease in salt concentration.[26] Physiological salt concentrations are approximately $0.15M$, with Potassium (K^+), Sodium (Na^+) and Magnesium (Mg^{2+}) the major biological contributors.[29]

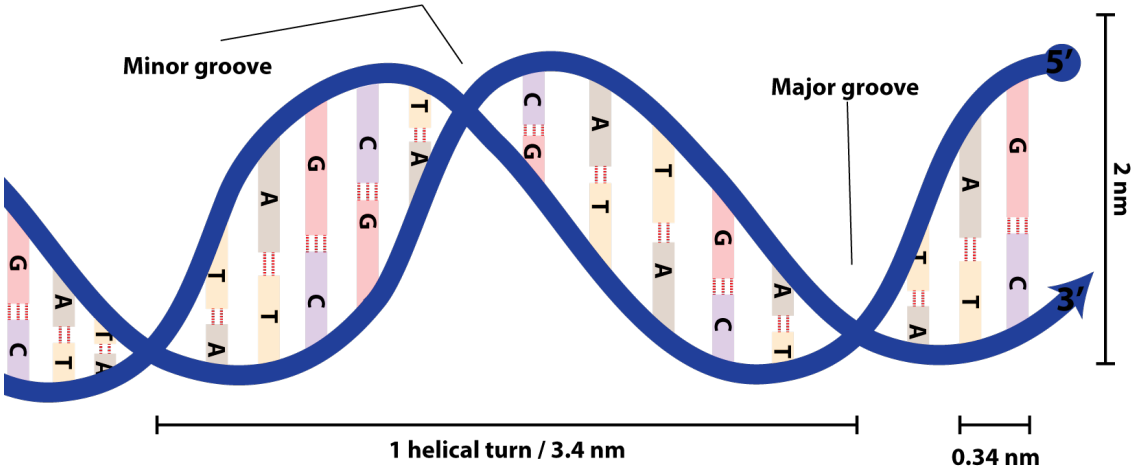

Figure 2.4: ***A schematic diagram depicting the helical structure of duplex DNA.*** *The two strands of DNA wrap around one another, displaced from the central axis of the polymer. This gives rise to the occurrence of minor and major periodic grooves.*

When considering physical properties, the native form of DNA has a helical pitch of approximately 3.4 nm with the separation between each base pair constituting 3.4 Å. This constitutes approximately 10.4 base-pairs per turn and is commonly referred to as B-form DNA. This form of DNA is found to be the native morphology adopted within a physiological environment, although it is important to note that other forms do exist, typically transiently within a cell or *in vitro*. It has been shown that this B-form has an approximate diameter of 2 nm, with a Hydrogen bond separation of 2 Å for both A-T and G-C base-pairs.[30]

Within the double helical structure, the bases are all stacked within parallel planes due to the hydrophobicity of their planar faces. This stacking is defined by the interaction of adjacent π orbitals and electrostatic interactions between neighbouring bases. Thus the centre of the double helix excludes water entirely, which increases stability in its native aqueous environment.

Interactions between sequential heterocyclic rings of the bases causes a knock on effect in the dipole characteristics of neighbouring nucleotides, in this way the stacked bases are stabilised by the London dispersion force.[31] Furthermore, due to the heterocyclic structure and hetero-sequential nature of nucleic acid bases arranged in DNA, the close interaction energies of charged planar faces is minimised by offsetting the nucleotide rings from one another, giving rise to the helical pitch described above (figure 2.4).[31] Additional contributions to helical conformation are instigated by the pucker of the de-oxyribose sugars.

Moreover, due to the energy-minimised structure of B form DNA it has favourable mechanical characteristics. Most interesting is its rigidity, with an accepted persistence length of approximately 50 nm when in aqueous solution. Thus DNA is considered a semi-flexible polymer with respect to the Kratky-Porod worm-like polymer chain model.[32]

2.2 Understanding Proteins

Proteins may be considered the most functionally complex and versatile of all biological molecules, responsible for the majority of catalytic, structural and motor functions in the cell. As such, they are found to make up the majority of cellular dry mass. The basic subunit of a protein is the amino acid, which are linked together to form linear polymers known as peptides. The large diversity of protein function is dictated by the complex three dimensional folding of these amino acid polymers, across four levels of organisation.

This versatility allows proteins to adopt a huge variety of roles within a cell, such as signal transducers, cargo carriers, ion channels and pumps, antibodies, toxins, hormones and even sources of bio-luminescence.[23]

A subset of proteins, called enzymes, make particular use of the three dimensional intricacy of their surfaces to perform the catalytic function of promoting critical chemical reactions. Such enzymes enable critical biochemical reactions at a rate suitable for life to occur. In particular, the interaction of proteins with nucleic acid forms a central set of process in biology. The processing of DNA is of critical importance to all life, from unwinding and untangling; to replication; to

recombination and maintenance.[23] It is the proteins involved with this latter function that are of interest to bionanotechnology and this project. In the following section, the biochemistry and organisation of proteins will be described.

2.2.1 Amino Acid Chemistry

There are 20 types of amino acids utilised as the constituent building blocks of proteins in biological systems.[23] The function of the final protein is dictated by its folded structure, which is in turn dictated by the unique sequence of amino acids. Thus, it is important to discuss the chemistry of the subunit in order to understand the functional versatility of proteins.

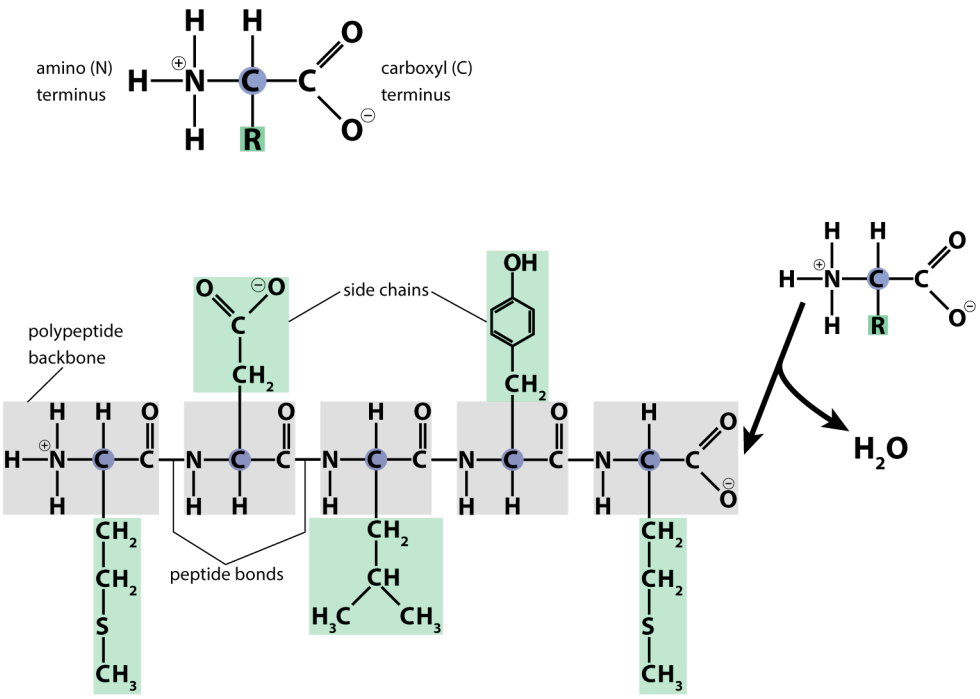

Figure 2.5: ***The chemical structure of amino acids and peptides.*** ***Top*** *The basic chemical structure of an amino acid is shown. The amino terminus (N), alpha carbon (blue), functional group (R, green) and carboxyl terminus (C) are indicated.* ***Bottom*** *Amino acids are joined through peptide bonds, releasing water. The resulting peptide has an inherent polarity due to the asymmetry of its terminal amino and carboxyl groups. The amino acid functional groups protrude perpendicular to the peptide backbone.*

The basic chemical structure of an amino acid can be seen in figure 2.5. Amino acids contain an integral polarity centred around an α carbon, with an amino group (N terminal) and a carboxyl group (C terminal) which are both ionised at neutral pH. The molecular mass of an amino acid is measured in Dalton (Da) which is equivalent to one atomic mass unit (u), such as a single nucleon or $1 gmol^{-1}$. The average molecular mass of an amino acid is found to be 110 Da.[23] Interestingly

all amino acids exist as stereo-isomers, L and D, due to the asymmetry of the α carbon. It is important to note that all proteins are formed from L-amino acids exclusively.

This charged polarity enables the formation of an amide linkage between the N and C termini of neighbouring amino acids, known as a peptide bond (figure 2.5). The addition of amino acids through sequential peptide bonds forms a linear polymer, typically described from the N to C terminus (left to right) and commonly referred to as a polypeptide.

The differential chemical functionality of amino acids is derived from the side chains that protrude perpendicular to the polypeptide backbone, denoted by R in figure 2.5. Amino acids are grouped based on the chemical functionality of their side chains; acidic; basic; uncharged polar; and non-polar.

The functionality of a protein is known to be a consequence of its polypeptide – referred to as primary – sequence, with the average length of a protein being 300 amino acids. Since there are 20 possible amino acids, then the possible combinations scale as 20^n for chain length n. In other words, for the average protein, there are a possible 20^{300} different polypeptide chains. This is an immense amount of possible sequence space, that would consume all the atoms in the universe in order to sample simultaneously. However, of these possibilities only a small fraction are capable of adopting a stable three dimensional structure, which have been carefully selected for through evolution based on the efficiency of their function. As will be examined in the following section, this three dimensional structure is formed from the combination of secondary structural domains and governs a proteins functionality.[23]

2.2.2 Intermolecular Bonds and Secondary Structure

The folding landscape of a polypeptide chain is limited by the possible bond angles available to it. There are three bonds from each amino acid subunit that contributes to the polypeptide chain. The peptide bond may be considered as a rigid planar unit, unable to rotate about the carbon - nitrogen bond and is thus fixed. In contrast, the singular bonds between the amino and carboxyl groups and the α carbon are very flexible. By convention the dihedral angle of rotations are referred to as psi (ψ) and phi (ϕ) for the $C_\alpha - C$ and the $N - C_\alpha$, respectively. Due to steric hindrances between atoms within an amino acid there are only a few pairs of ϕ and ψ angles permitted, which can be described by a Ramachandran plot. The permitted bond pairs are shown to fall largely into two constituent groups, which describe the most common secondary structural motifs; the α helix and β sheet.[33]

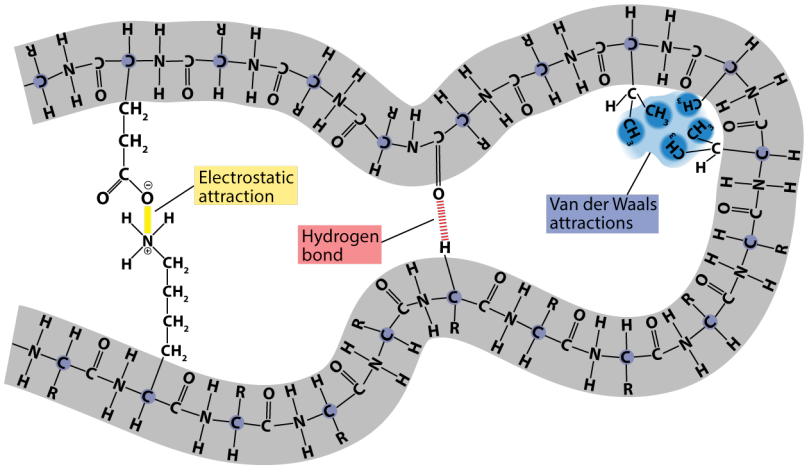

Figure 2.6: ***The formation of protein secondary structure.*** *Protein secondary structure is formed through the interaction of the amino acid functional groups through; electrostatic attraction (yellow), hydrogen bond formation (red) and Van der Waals attraction (blue). Additional formation forces come from the hydrophobic exclusion of the surrounding aqueous media.*

The folding of a nascent protein chain is guided by many weak non-covalent bonds that form between the side chains and the polypeptide backbone. These include; hydrogen bonds; electrostatic attractions; and Van der Waals attractions which can be up to 120 times weaker than a covalent bond. Despite this, the sum of all the weak interactions allows the formation of stable structural domains (figure 2.6).[34,23]

Additional stability occurs as a consequence of the hydrophobic effect, where the non-polar side chains are forced together in an aqueous environment. This occurs in order to reduce the disruption of the hydrogen bonding network of the surrounding water. As a result the non-polar residues have a tendency to be buried within the folded protein structure and the polar groups exist on the outer surface, exposed to the aqueous environment. Thus the distribution of the polar and non-polar amino acids is of critical importance to the stabilisation of the secondary domains. As a consequence it can be considered that the final folded structure of a protein represents its free energy minima.[34,23]

Proteins increase in geometrical complexity through the formation of independent structural domains, known as secondary structure. Although the folding pattern of each protein is unique, there are two commonly repeated domains; the α helix and the β sheet (figure 2.7).[35] The prevalence of these structures is in part due to the permitted pairs of ϕ and ψ angles and the H bonding between the N —— H and C ═══ O of the polypeptide backbone.

β sheets (figure 2.7) can exist in parallel - with chains running in the same direction - and anti-

parallel - with chains that fold back upon themselves, running in alternating directions. Both types of β sheets are found to be very rigid, due to the large number of hydrogen bonds that cross link the neighbouring polypeptide chains perpendicular to the peptide bonds. As a consequence of its structural integrity, this motif is commonly found at the core of many proteins.[35,23]

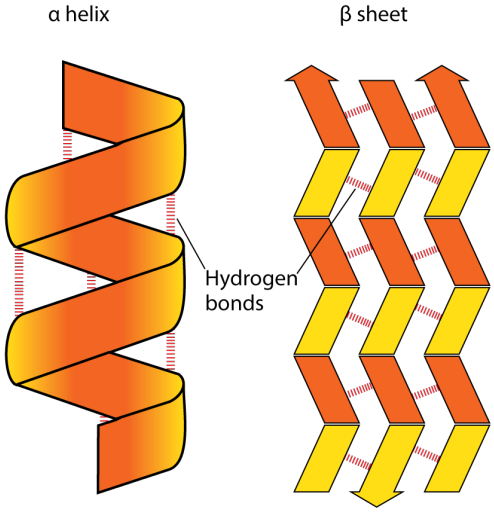

Figure 2.7: ***A schematic representation of α helix β sheet structure.*** *These common secondary structures form due to specific hydrogen bonding (red). The direction of the peptide chains traversing the structure is indicated.*

α helices (figure 2.7) are again formed from hydrogen bonding between the C$=$O and N$-$H in neighbouring amino acids. These form within a single polypeptide chain which twists around itself to form a right handed helix containing 3.6 amino acids per turn. Typically this results in the exposure of the hydrophilic polypeptide backbone, shielding hydrophobic residues within the core. Alternatively, where hydrophobic non-polar residues are exposed upon the external surface of the α helix, these peptides typically reside within the hydrophobic environment of a lipid membrane. Where these hydrophobic (non-polar) side chains exist uniformly on the same side, such that they rise parallel along the polypeptide axis, then multiple helices - typically two - are able to twist around one and other in order to shield the hydrophobic residues within. This structure is known as a coiled-coil, which is found in many elongated rod-like proteins.[35,23]

2.2.3 Tertiary and Quaternary Organisation

The final functional state of a protein is orchestrated by the association of multiple secondary structures to form a tertiary structure. This level of organisation maybe considered the final structure in many cases.

Within this tertiary level, structural and evolutionary studies have demonstrated the importance of separate modular subunits or domains that are responsible for the individual process of an overall protein function. For example, in the case of the restriction enzymes, one domain will be

responsible for sequence recognition and another domain is responsible for cleavage of the nucleic acid backbone.[23]

These subunits consist of small regions of polypeptide chains including α helices and β sheets, up to 350 amino acids in length (depending upon the overall protein size). Evidence from structural studies demonstrates that the polypeptide chain entirely forms one domain prior to crossing into and folding the next domain.[36] Domains are often conserved throughout evolution, resulting in the classification of protein families.

In the same way that small regions of a single polypeptide chain can form discrete functional domains, fully folded proteins are capable of interacting to form complex functional structures. This arrangement is described as a proteins quaternary structure (figure 2.8).

The same non-covalent bonds, described earlier (section 2.2.2), are responsible for the formation of tertiary (through intra-molecular bonding) and quaternary (through inter-molecular bonding) structure. Where this occurs between identical proteins forming a symmetric complex, this is known as a dimer and the constituents are referred to as protein subunits (figure 2.8). The formation of quaternary protein complexes is not restricted to identical proteins, but may exist as multi-subunits.[23] For example; Haemoglobin, consists of two α-globin and two β-globin in a large symmetrical ring arrangement.[37]

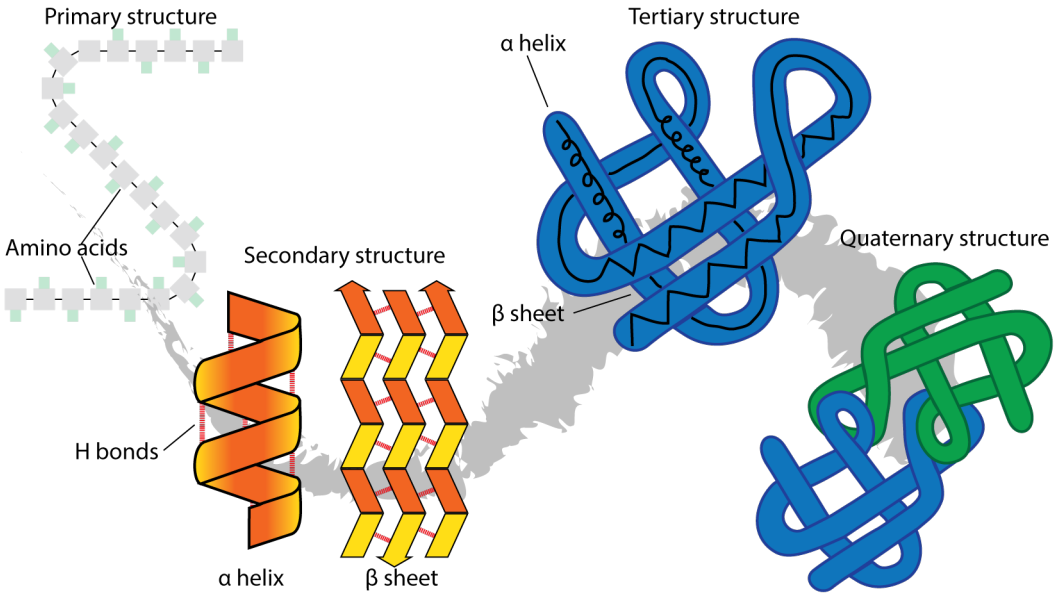

Figure 2.8: ***A summary of protein structure hierarchy.*** *Protein structure is formed from the hierarchical folding of peptide sequences to form functional domains and structures (left to right).*

2.2.4 Conclusions

This chapter has examined the structure and function of DNA and proteins, from their constituent chemistry to their functional conformations and how these relate to their native functions.

As described, both molecules are formed from the linear polymerisation of monomeric subunits, the sequence of which dictates further hierarchical assembly. Natively, this results in a linear duplex polymer for DNA. However, in proteins this sequence results in a three dimensional folding due to the cumulative effect of many intra-molecular forces, forming a functional unit.

Furthermore, proteins are not restricted to a rigid conformation. They commonly have regions of conformational flexibility that are linked to the orchestration of specific chemical events. Typically, these are found to be robust even in non-native environments. As such many proteins have been shown to be able to catalyse specific reactions *ex vivo* which has enabled the development of many biochemical assays and tools for the manipulation of DNA in molecular biology.

The adoption of a host of enzymes has given biologists a toolbox with which DNA can be cut, copied and pasted - using restriction enzymes, DNA polymerase and DNA ligase, respectively - to create *de novo* genetic arrangements. Furthermore, given the detailed understanding of how protein structure relates to function, it has been possible to add or remove specific protein domains to provide further or more specific functions compared to their native counterparts.

From this chapter, it is clear that a deep understanding of the intrinsic properties of biomolecules is critical in guiding how they may be harnessed for artificial applications. In the next chapter, how the sequence structure of DNA can be simply manipulated to form novel artificial structures and functional devices is introduced. In addition, further hierarchical assembly on DNA structures is explored through the exploitation of recombination enzymes, specifically RecA.

Chapter 3

Literature Review

In this chapter the extant literature regarding the development of DNA nanotechnology, its emerging applications and the potential use of RecA for spatially addressing these structures to provide heterogeneous assemblies is examined. In the first instance, the theory underlying the use of DNA as a construction material and the development of DNA nanotechnology is discussed. This section identifies the principles that govern the structure of DNA and how manipulation of these factors can lend themselves to the use of DNA as a robust construction material. Following this, key developments in the field are highlighted, including the development of woven motifs, the inception of scaffolded origami methodologies and the use of ssDNA "brick" systems are examined. Finally a discussion of the arising applications for DNA-organised systems, including nano photonic and electronic systems, platforms for studying biological molecules and emerging medicinal devices is presented.

In the second half of this chapter the homologous recombination protein, RecA, is reviewed in relation to its use for spatially addressing DNA nano-architectures. In order to do this, the current understanding of the structure of the RecA protein and its polymer is first described. Subsequently the homologous recombination function of RecA is examined, from the formation of the active nucleoprotein filament, the alignment of sequence homology and final strand exchange. This is followed by a description of the extant examples of the use of RecA in nanotechnology applications.

In the latter part of this chapter, the current understanding of the homology searching mechanisms of RecA are explored with relation to the methods of study. This is used to demonstrate the sometimes conflicting observations that arise from the ensemble or indirect approaches used to date. A contrast is, therefore, drawn between the approaches adopted by others and the direct

single molecule approach undertaken in this thesis. With respect to this, the applicability of the AFM - in particular the realisation of high speed AFM - is highlighted with examples of previously studied biomolecules, in order to justify its selection for this study.

3.1 DNA Nanotechnology as a Template for Bottom-up Hierarchical Assembly Regimes

The concept of structural DNA nanotechnology (SDN) was first envisaged by Nadrian Seeman in 1980[1], as a method for creating periodic arrays within which proteins could be hosted in specific orientations for crystallographic studies. Seeman's original inspiration was the M.C. Escher woodcut "Depth", which he postulated resembled a set of branched DNA junctions. He went on to establish a base set of branched junctions and further constructed a 3D cube from the complex interweave of several DNA strands as a proof of principle (figure 3.1 D).[1]

This inspired vision of utilising DNA as a material for the rational design of self-assembling nanostructures spawned the entire field of DNA nanotechnology. This field has advanced rapidly over the proceeding 30 years, to include a vast array of structural, functional and computational systems produced from DNA.

Underpinning this concept is the assembly of immobile branched junctions joined through the use of sticky end cohesions, enabling the growth of structures in multiple dimensions through iterative addition of junction units.[38] With respect to this, it becomes apparent that we must first discuss DNA from a material perspective in order to examine its exploitation for construction.

3.1.1 DNA as a Structural Element

DNA has been found to be an ideal biomolecule for the construction of nano-architectures due to its: abundance and availability; chemical addressability; robust and predictable helical geometry; and the specificity of the interactions through which it self-assembles.

In nature, DNA is the carrier of genetic information, based upon the specificity of the Watson-Crick base pairing regime,[24] as described in section 2.1. In principle, structural DNA nanotechnology utilises the fidelity of this molecular recognition to design the specific interaction geometry of individual DNA strands in order to form artificial structures.[1,38]

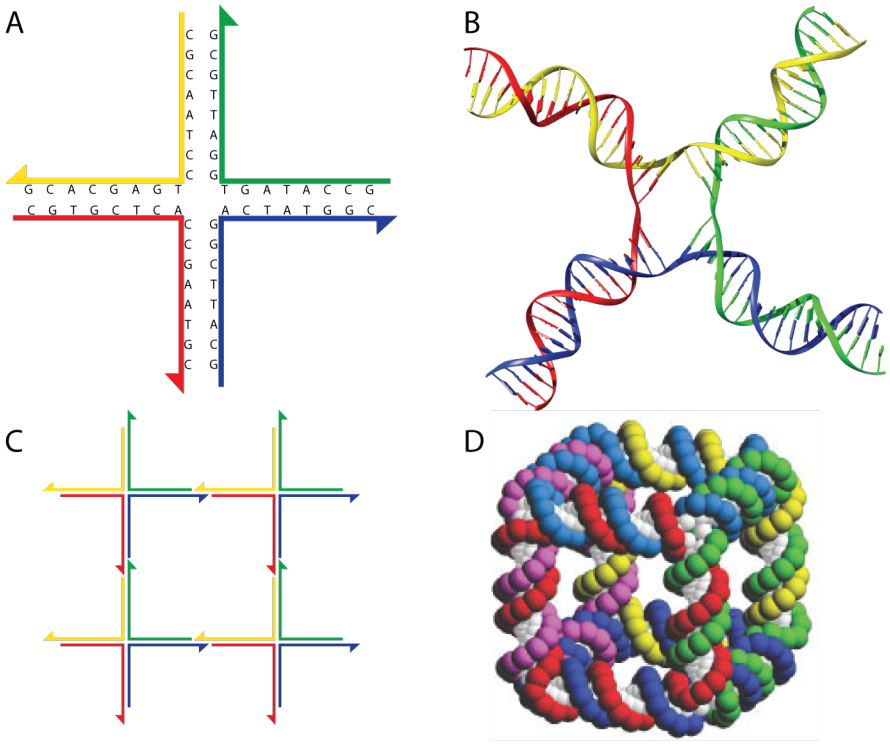

Figure 3.1: ***A set of schematic diagrams depicting the structure of four arm DNA branched junctions.*** *A four way junction formed from four partially complementary oligonucleotides with asymmetric sequence structure eliminating branch migration. Shown schematically* ***(A)*** *and as structural ribbon model* ***(B)****. Branched junctions can be sequentially joined to create periodic 2D lattices* ***(C)****. The same principles enable a cube to be created from six interwoven oligonucleotides* ***(D)****.* ***(B)*** *Generated using Chimera from PDB file: 3CRX.* ***(D)*** *Adapted from Seeman et al.*[1]

Although other base pairing paradigms exist (see section 2.1) which potentially open up the structural coding capacity of DNA further, the use of such would directly undermine the specificity of DNA hybridisation that thus far is the critical credential of nucleic acids as a construction material. Thus sequence designs are limited to the canonical AT and GC base pairing, with the differences in thermodynamic stability of the 2 and 3 bonded bases, respectively, a critical parameter in sequence determination.[1,30]

The basic unit of any DNA architecture is the branched junction, which is formed at the intersection of two non-canonical linear DNA duplexes, immediately expanding the structure in multiple dimensions. Here, we can imagine four strands of partially complementary single stranded DNA, such that half of every strand is complementary to half of another. Thus, the pairing of all complementary partners creates four individual duplexes linked at a single branch point (figure 3.1 A).

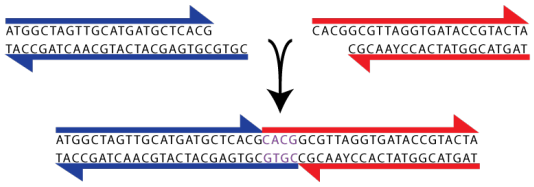

Figure 3.2: A schematic diagram depicting the hybridisation of DNA overhangs to form a single structure.

In nature these appear in the form of transient branched structures known as a Holliday junctions which are found as the intermediates of DNA recombination events (figure 3.1 B).[39,40,41] These structures are resolved enzymatically through the migration of the branch point along symmetrically sequenced arms, forming two separate DNA duplexes.[41,42]

Structural DNA nanotechnology adapts these structures by removing the sequence symmetry around the branch point and hence ceasing the migration of the joint (figure 3.1 A). Natively, Holliday junctions exist as four way junctions, however this can be expanded to in excess of twelve arms for nanotechnology purposes. In practice, increasing branch complexity is difficult to achieve due to limitations in the available unique sequence, given an alphabet of only 4 nucleotides. However the introduction of a 4 arm branch remains energetically favourable compared to that of the four separate ssDNA constituents free in solution – at approximately $+ 1.1\ kcal\ mol^{-1}$.[42]

It is important to consider that due to natural fluctuations within the environment, all DNA structures "breathe" with the H bond distance fluctuating around 2Å. This is particularly acute at branch points between multiple DNA strands - where the base-pairing interactions are typically most stressed - and as a consequence may lead to the dissociation of the structure. To counter this, vulnerable regions are typically strengthened through the inclusion of GC sequence repeats, known to confer the highest T_m (section 2.1.2) (figure 3.1 A).[30,26]

The further expansion of junctions to form higher order structures is dependant upon the complementarity of terminal single stranded regions known as "sticky ends" (figure 3.1 C).[30] These elements arise in nature following the action of restriction enzymes, allowing for the cut and paste mechanisms common within molecular biology. Where sequence complementarity exists, hybridisation is able to occur regardless of whether it exist within the same canonical species, resulting in a continuation of the helical geometry (figure 3.2).

The predictability of this geometry allows for very precise control over the orientation of subsequent attachments[30] without the need to continually asses the relative positions of species at each step of construction. It is important to consider that the T_m of such interconnections must be higher than the individual species that make up the subunit itself, thus inferring global stability as the structure grows.

3.1.2 Development of Motif Structure

Building on this, a host of two and three dimensional lattice structures based on a small set of arrayed junctions, were demonstrated.[43,44,45,46,47,48,49,50] In contrast, at the same time and utilising the same concepts, multiple discrete polyhedral constructions were demonstrated.[51,52,53,54,55,56,57]

Of particular note, the work by He et al, demonstrated the development of large supramolecular polyhedra.[53] Fascinatingly in this example, only three DNA species were used to produce a three point star motif that was utilised to form a large range of polyhedra by varying the concentration of the final DNA motif available for subsequent assembly.[53] The structures formed, ranged from a basic tetrahedron to a Bucky ball (Buckminsterfullerene), at a DNA concentration of 500 nM. The curvature of such structures was created due to the clever design of the helical twist, separating the joint by incomplete helical turns, inferring global curvature. Where the curvature is accumulated on the same face of the structure construction in three dimensions is favoured. Principally it was noted that working at nano-molar concentrations favours discrete 3D formations compared to micro-molar concentrations favouring large 2D lattice structures.[53] Indeed this approach has later been applied to construct tubular structures for templating nanowires.[58]

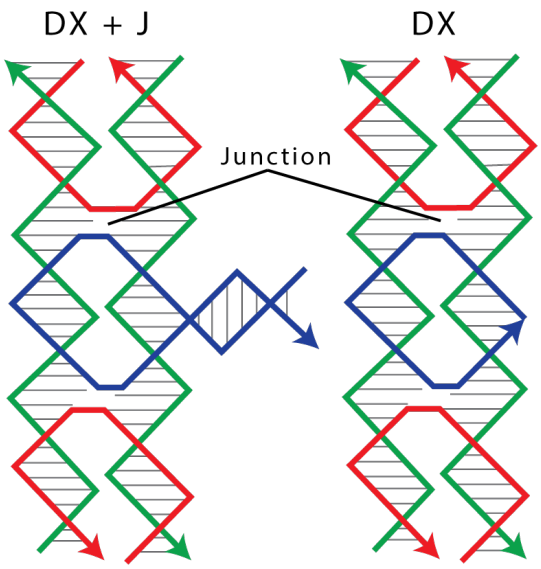

Figure 3.3: A schematic diagrams depicting the structure of a double crossover (DX) and double crossover junction (DX+J). Different DNA species are indicated by colour. Reproduced from Seeman et al.[38]

Central to achieving this eruption of successful structures, was the development of a set of core structural motifs. These were designed to overcome the poor mechanical rigidity of dsDNA over long length-scales 2.1.2), which is described as a semi-flexible polymer.[32]

This is an important consideration, as critically it lacks the robustness required to form rigid lattices in two dimensions and certainly was found to cause problems when considering three-dimensional structures.

Work by Winfree et al combated this problem by weaving parallel DNA duplexes together

increasing global rigidity by limiting the gyration of the molecules about one and other. This approach is analogous to the braiding of rope.[43,39,59] In its simplest form, a double crossover (DX) motif can be constructed from two parallel DNA duplexes (figure 3.3). Here, each strand contributes to either DNA molecule, associating the two duplexes together at the exchange points. A robust DX motif is created where two crossovers occur separated along the helical axis between two parallel duplexes (figure 3.3). This can be further extended to include a third DNA duplex in parallel to create a TX motif or set perpendicular to create a junction or DX + J motif. It is interesting to note that although crossovers can be carried out between strands of both polarities, exchanges between strands of opposing polarities – i.e. where the ssDNA molecule reverses direction at the crossover – confer the most predictable topologies.[38]

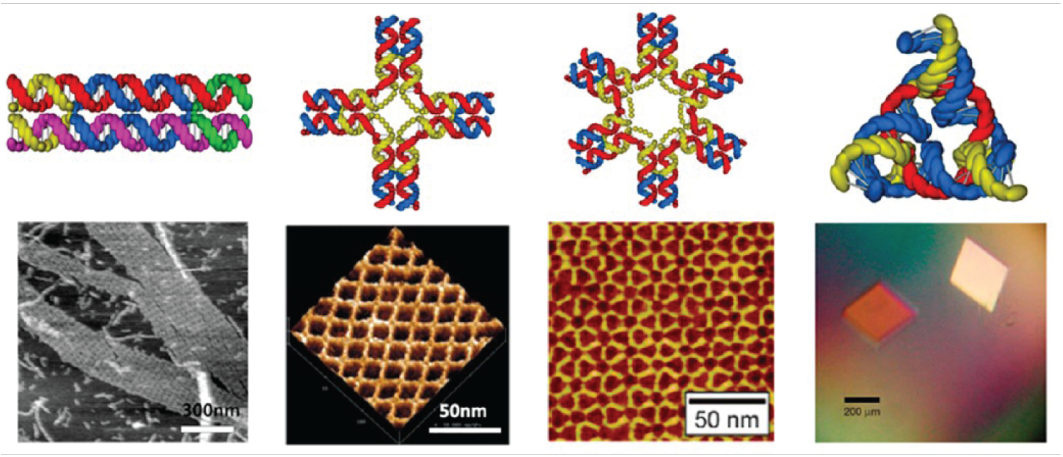

Figure 3.4: ***DNA nanostructure motifs and their associated 2D arrays and 3D crystal. (Top)*** *schematic helical arrangements and* ***(bottom)*** *AFM images of assembled 2D arrays and 3D crystal. Adapted from Zhang et al.*[60]

At this point, a distinction must be made between the original junction based arrays and the latter crossover based arrays (figure 3.4). It can be considered that by weaving multiple DNA strands together with multiple crossovers – e.g. double (DX) or triple (TX) – that small discrete units or "tiles" are created. Thus, these small tiles can then be tessellated and assembled together using short connecting species (figure 3.4, left).[61]

Further adaptation of these motifs was achieved by incorporating them within junctions, such as the three armed junctions created by Ding et al.[62] This DX triangle motif was created by combining a double crossover (DX) motif with a bulged triangle design, in order to overcome the flexibility of previous triangular junctures.

By arranging the three sets of parallel strands overlapping one another around the circumference

of the junction, a tensegrity triangle design was created. This has been shown to be one of the most robust motifs created to date, enabling the development of up to macroscale polyhexagonal lattices.[62,63] It is important to note that not only is the triangular junction itself very robust, but through the inclusion of DX motifs in all three arms, neighbouring junctions are joined through parallel sticky ends strengthening these traditionally flexible regions and preventing any flexing out from the growth plane. 3D lattices built with these tensegrity triangle unit cells have brought Seeman's original vision of orientated crystallography leaps and bounds closer.[46,1]

3.1.3 DNA Origami

In contrast to hierarchical motif construction, seminal work by Paul Rothemund changed the landscape for structural DNA nanotechnology, through the introduction of DNA origami.[2] This method directs the folding of a large ssDNA "scaffold" - typically a viral genome - in to a desired shape using short ssDNA oligonucleotides "staples" to pin distant regions together (figure 3.5A - E).

The use of an underlying scaffold strand circumnavigates the stoichiometric and purification problems associated with construction from many short species and allowed for near-quantitative yields of 2D tiles.[60]

To demonstrate this concept, Rothemund folded M13 viral DNA into a multitude of intricate shapes, including a "smiley face" which can be seen in figure 3.5 C - E.[2] Furthermore, in this seminal work Rothemund already eluded towards three dimensional construction (figure 3.5 G) and expansion of geometry via the tessellation of tiles (figure 3.5 H). This work suggested the notion of DNA origami tiles as molecular pin boards upon which protruding DNA loops could be incorporated, thus using a height difference to creating effective "pixels". Demonstration of this can be seen in figure 3.5 G where Rothemund wrote the letters "D", "N" and "A" on a tile.[2]

Despite the different approaches taken by Rothemund and Seeman - a comparison of which is given in figure 3.6 - they both inherently rely on the interconnection of multiple DNA duplexes through branched junctions. Indeed, the woven strength of DNA origami comes as a result of the incorporation of double crossovers between adjacent DNA duplexes throughout the structure (figure 3.5 F). Note that in origami designs, both the scaffold and the staple strands can contribute crossovers and individual staples can be involved in crossovers between multiple duplexes sequentially, depending on length.[2,64,65]

This methodology was immediately adopted by other groups and quickly expanded into 3D,

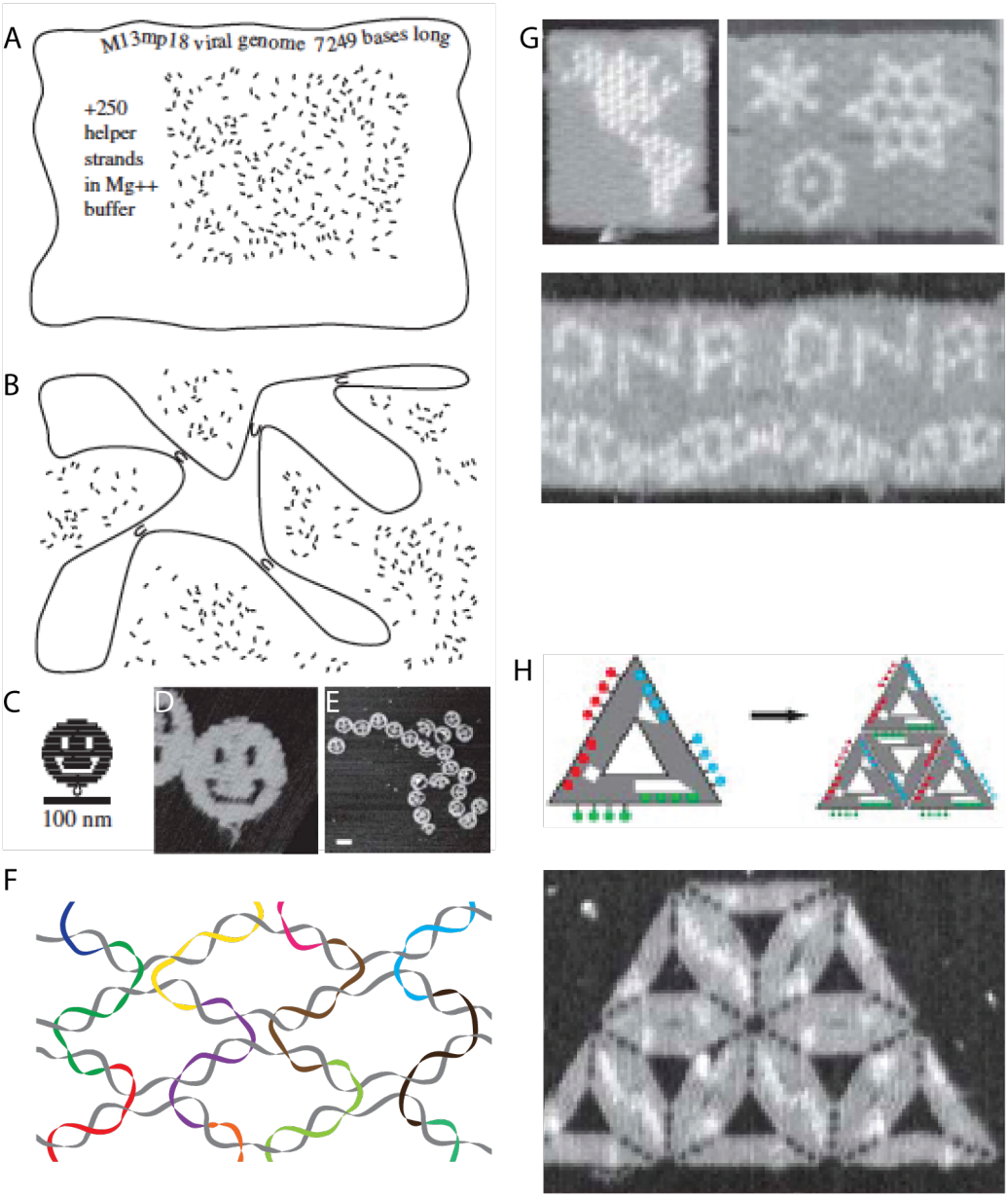

Figure 3.5: ***An overview of the DNA origami method.*** ***(A)*** *A large single stranded "scaffold" DNA molecule is mixed with a large number of short single stranded "staple" oligonucleotides.* ***(B)*** *Staples are partially complementary to distal regions of the scaffold, bring these regions together upon binding. Arbitrary patterns can be folded with high efficiency, such as a smiley face; shown schematically* ***(C)*** *and in AFM images* ***(D) & (E).*** ***(F)*** *The modelled internal weave of DNA strands within a folded origami structure is depicted schematically. The scaffold strand is shown in grey traversing horizontally through the structure. The staple strands are shown in a multitude of colours contributing to multiple DNA duplexes as they traverse vertically through the structure. Examples of 3D pin boards* ***(G)*** *and 2D arrays* ***(H)*** *of origami are depicted. Adapted from Rothemund.*[64]

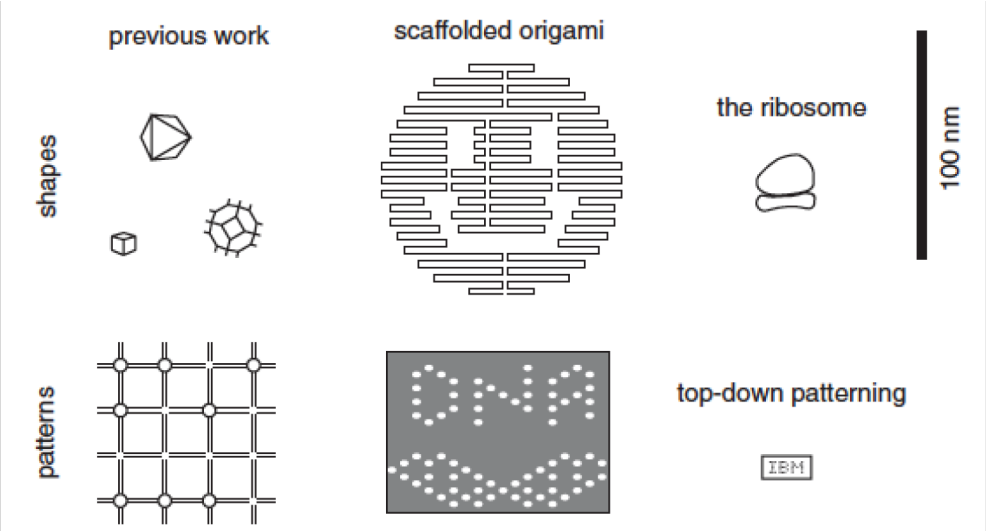

Figure 3.6: ***A size comparison of DNA nanotechnology types.*** *A schematic diagram comparing the relative sizes of junction based polyhedra and lattice structures (Left) with scaffolded origami structures (middle). For comparison, the sizes of a Ribosome and the state of the art micro-machining are depicted (right). Adapted from Rothemund.*[2]

through either the consecutive layering of sheets creating close-packing pseudo-solid objects,[66] or the joining of vertices to create intentionally hollow shapes with each face produced from a 2D origami tile.[67,68,69]

Meanwhile, other groups examined the growth of 2D tile networks by: scaffolded templating individual tiles;[70] base stacking at the termini of the helices;[71] or specific base pairing (akin to earlier DNA nanotechnology junction work).[65] Contrasting work has examined the expansion of origami tiles themselves, through the use of longer scaffold strands, with up to 51 kbp sequences demonstrated by Marchi et al.[72]

Considering 3D close-packed objects, DNA helices can be assembled together into either a square[73] or hexagonal lattice structure (figure 3.7).[66] This choice of packing lattice directly relates to the spacing of cross over points along the helical direction. With regards to the hexagonal lattice, crossovers spaced every 7 bp gives a rotation of connection points by 240° connecting the 3 neighbouring strands successively along the duplex. In this way, any particular pair of duplexes is connected twice - a double crossover - every 21 bp (figure 3.7). This approach follows the B-form helical pitch throughout the entire woven structure, but results in a porous weave of parallel helices.

In contrast, by weaving DNA in to a square lattice, each duplex connects to 4 neighbours, requiring the crossover positions to be placed every 8 bp by partially relaxing the helical pitch from 10.5 bp

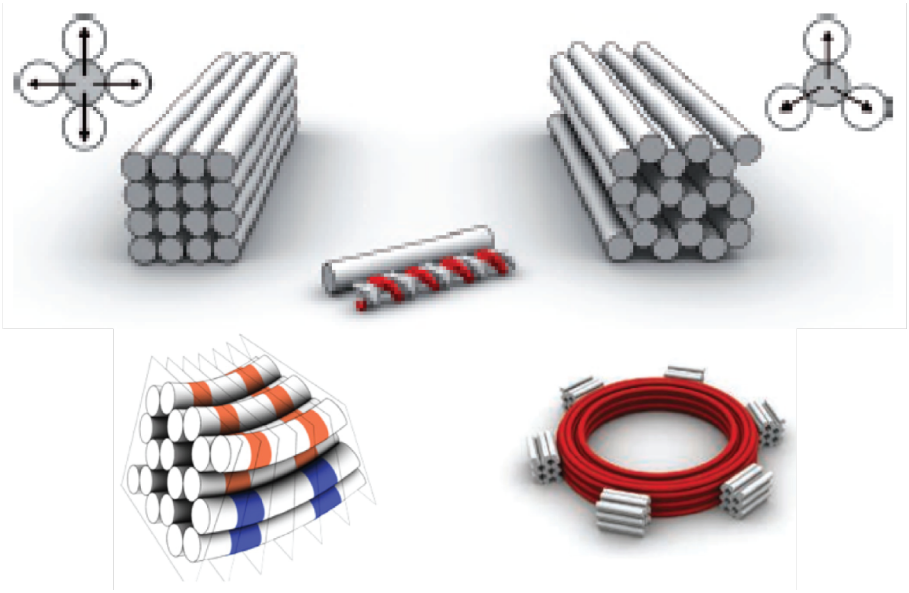

Figure 3.7: ***Origami packing geometries and induction of curvature.(Top)*** *DNA duplexes are packed onto either a square (left) or hexagonal (right) lattice.* ***(Bottom)*** *Global curvature can be induced into the structure by offsetting the crossover points between adjacent DNA duplexes (orange & blue). This enables construction of curved structures, such as gears. Adapted from Castro et al*[65]

to 10.67 bp per turn.[65] This inevitably adds a degree of global strain or twist throughout the planes of DNA duplexes, but results in a far closer packing of parallel strands - which is important when considering some of the applications explored later (section 3.1.4).

Moreover, the density and positioning of crossovers is crucial to dictating the global twist and curvature of the finalised design, where subtle deviations from the desired crossover positioning can cause intentional under-twist, over-twist or axial strain.[74,75,65] When considering Rothemunds original single layer tiles produced with crossovers spaced a standard 16 bp apart between each pair of duplexes, these undoubtedly contain a global curvature in solution. However, this is largely neutralised through electrostatic binding to a support surface, such as those used for microscopy observations.[65,2]

In addition to the flurry of structures produced with DNA origami method, other groups have expanded the horizons further, through reductive folding approaches, attempting to either remove the scaffold or staples.

Work by Geary et al has expanded the origami approach to RNA, which due to its more flexible backbone, demonstrates motifs unavailable to DNA, such as the “kissing loop”. This system was designed to be a single long RNA molecule, akin to a biological mRNA (messenger RNA) that contained self complementarity. Thus, RNA tiles were designed to fold in the absence of additional

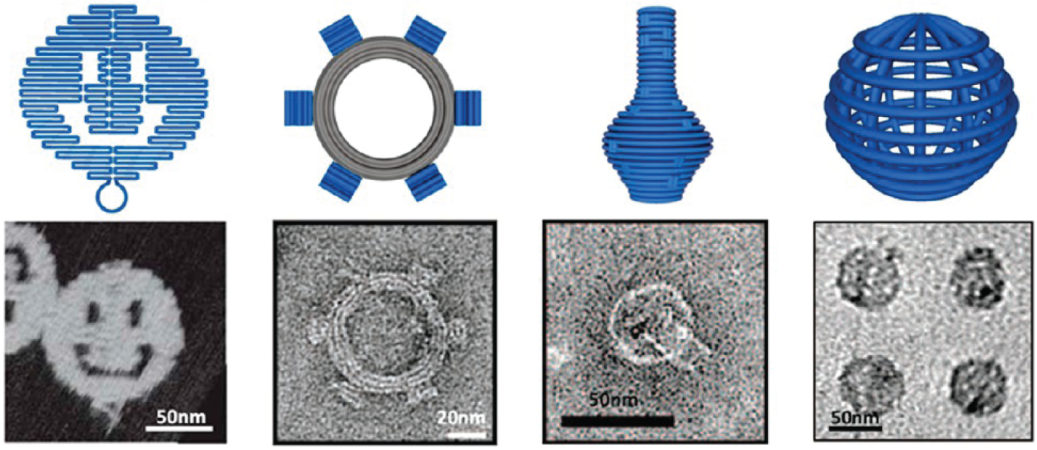

Figure 3.8: ***A Summary of origami of 2D and 3D origami structures demonstrated to date. (Top)*** *schematic diagrams depicting the routing of the scaffold strand.* ***(Bottom)*** *AFM (smiley face only) and TEM images depicting the described structures. Adapted from Zhang et al*[60]

staple strands and hence, this technique is termed "single stranded origami".[76]

This system was based upon the known intricate RNA secondary structures found throughout biology, all of which are produced through transcription from a DNA template. Amazingly, Geary and co-workers were able to replicate this process *in vitro*, folding their RNA tiles co-transcriptionally as they were produced from a synthetic plasmid using a RNA polymerase.[76]

An alternative approach was demonstrated by Yin et al, where a set of interlocking ssDNA sequences that form localised connections with one another were developed. These so-called "bricks" contain 4 domains which are complementary to 4 different bricks and are designed to create 2D or 3D canvases, in the absence of a scaffold strand. From this blank canvas, the desired design is "engraved" by selectively omitting or including specific bricks in a one pot folding process. This approach is inherently automatable, with each "brick" representing a pixel (in 2D) or a voxel (in 3D), a pipetting robot can easily be programmed to include only the necessary components for a particular design. Hence, Yin and co-workers rapidly demonstrated a vast number of different designs in a very short time - including a full 26 letter alphabet - highlighting the versatility of the approach[77,78,79]

More recent developments have focussed upon merging the scaffold driven origami approaches with the earlier junction based polyhedra methods to develop novel routing algorithms for wire-framed structures. These approaches have enabled the researchers to generate novel, non-symmetrical and arbitrary wire-frame structures that would be incredibly difficult via previous methods, thus further expanding the versatility of DNA nanotechnology (figure 3.9).[80,81,82,83]

3.1.4 DNA Nanotechnology Applications

As with anything youthful, the first stages of development are involved with investigating the abilities and pushing the limits. However, as time passes DNA nanotechnology must inevitably begin establishing its practical approach and applications.[60]

The power of DNA nanotechnology in real-world application is the accurate spatial and heterogeneous templating of external elements to form practical devices including: circuits; actuators; and sensors at the nanoscale.[60] Not only can final devices be templated using these methods, but they allow experimental platforms upon which specific properties of a material can be investigated.[84]

For example, the arrangement of metallic nanoparticles, quantum dots[85,86,87,88,89,90] and organic chromophores[91] have all been shown – enabling investigation into the spatial dependencies of these photonic elements.[92,93,94] The ability to accurately control the spatial arrangements of such elements is seen as a pathway to enable the development of phototransistors, optoelectronic devices, new generations of solar cells, laser diodes and light emitting diodes (figure 3.10).[60]

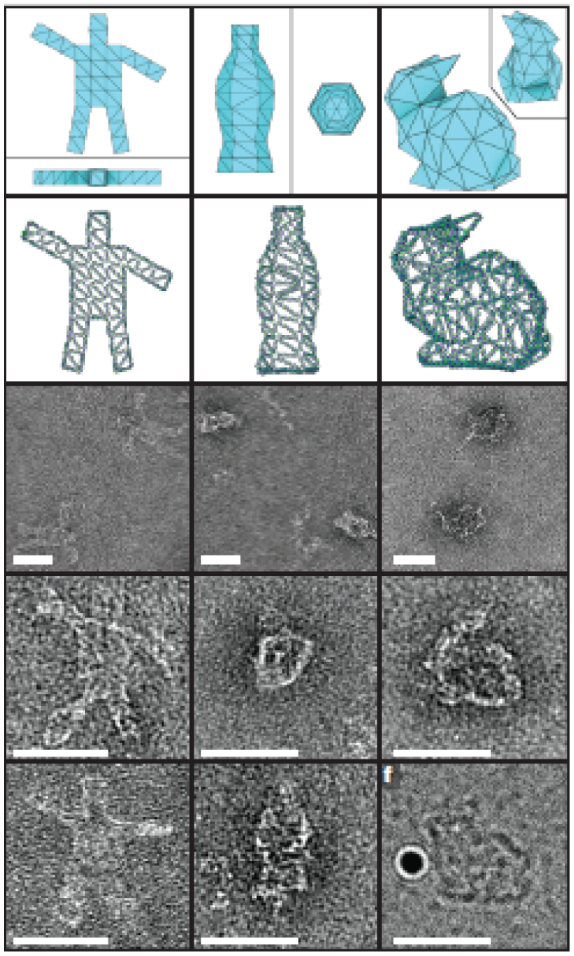

Figure 3.9: ***Scaffold routed DNA wire-frames.*** *Three example designs; a waving stick man, a bottle and the Stanford bunny are demonstrated. In columns (top - bottom); the initial wire-frame design, algorithmic scaffold route and cryo-electon microscopy images. All scale bars = 50 nm. Adapted from Benson et al.*[80]

One such example is the work by Liedl and co-workers, which demonstrated a staircase arrangement of gold nanoparticles templated around the circumference of a DNA nanotube. This arrangement enabled the specific tailoring of the surface plasmon resonance of these structures (figure 3.10).[87] Related work by Dutta et al, demonstrated a

similar arrangement of organic chromophores in order to direct a cascading energy transfer through an artificial light harvesting complex.[91] Alternatively, origami templates have been developed as lithography masks, in attempts to extend the reach of more traditional top-down construction approaches by combining them with bottom-up assemblies.[95,96,97,98]

Thus far this text has described the spatial arrangement of components using DNA templates. However, it is necessary to consider that any practical application requires tight control over the orientation, spatial distribution and interconnection of individual DNA-organised devices themselves. This has been a difficult problem to tackle, as DNA nanostructures are formed and typically functionalised in solution. As a result, deposition onto solid surfaces often results in randomised arrangements and orientations. This poses problems for characterisation as large areas must be mapped to find correctly deposited devices.[3] Furthermore, applications such as nanoelectronics that require the structures to be externally addressed or nanophotonics where specific spatial arrangements of nanodevices are required are hence drastically limited by current surface arrangement methodologies - or lack there of.[3]

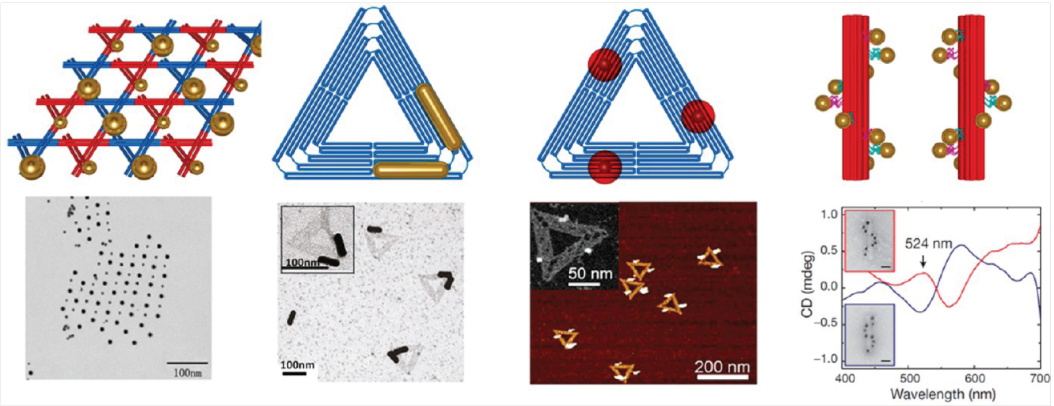

Figure 3.10: ***Representative examples of DNA-organised inorganic nano-particles.*** ***(Top)*** *Schematics depicting the DNA organisation of inorganic nano-particles, shown alongside corresponding AFM or TEM images. Adapted from Zhang et al.*[60]

In response to this, several efforts have been made to template the surface arrangement of DNA based devices through chemical,[99] lithographic[100,101,102,103] and topographic confinement.[104] Of particular note is the work by Gopinath and Rothemund[3], who built upon the lithographic placement methodology of Kershner et al.[100] Here, binding sites of identical shape to the origami intended for deposition were lithographically patterned through a passivated layer onto SiO_2 surfaces. O_2 plasma etching created silanols which ionise at the deposition pH to give a negatively charged binding site. Deposition of the origami causes preferential binding and alignment into these designed wells through the divalent cation bridging of the negative DNA backbone to the

silanols mediated by Mg^{2+}(figure 3.11).[3] The introduction of amine groups on the inner edge of the origami triangle, enabled the subsequent covalent coupling of the origami to the surface, locking the arrangements in place. This then enables the surface origami arrays to be exposed to non-native chemical environments for downstream processing or additional functionalisation.[3]

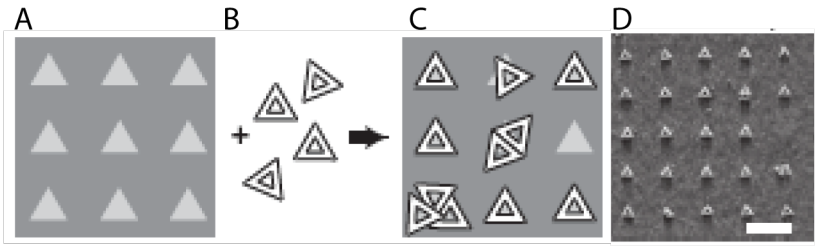

Figure 3.11: ***Surface organised DNA Origami tile arrays.*** ***(A)*** *A $Si0_2$ surface is plasma etched through a passivation layer, patterning triangular regions of negatively terminated silonls.* ***(B)*** *DNA origami in Mg^{2+} is deposited and arranges preferentially into the lithographically patterned binding sites* ***(C)****. An AFM image depicting an exemplar surface array produced with this method. Scale bars = 400 nm. Adapted from Gopinath et al.*[3]

Alternatively, with respect to biological applications, DNA scaffolds have been used to organise proteins for both study and device formation.[105] The Tuberfield group arranged proteins attached to periodic 2D tile arrays, enabling study of their structure through cryo-EM.[106] Similarly, work by Shih et al, isolated proteins within DNA nanotubes allowing the structural study of previously uncharacterised membrane proteins using nuclear magnetic resonance imaging (NMR).[107] Both of these examples approach Seeman's original vision of DNA nanotechnology as a host network for protein crystallography.[1]

DNA and RNA templates have also been used to confine multiple proteins in specific orientations and geometries in order to study metabolic processes,[60] signalling cascades,[108,109,106,110,111,112] the spatial dependencies of motor proteins,[113,114,115] single molecule force sensing platforms,[116] biosensing assays[117,118,119] and the observation of single molecule DNA or protein-DNA interaction events.[120,121,122,123,124,125,126,127,128]

In addition to the study of biological molecules, unique applications of DNA origami in medicine are being demonstrated, in particular as targeted local delivery systems or devices for triggering cellular signalling cascades.[60,129] For example, the last decade has seen the development of DNA-organised systems for the delivery of: the anticancer drug doxorubicin (DOX);[130,131,132,133]

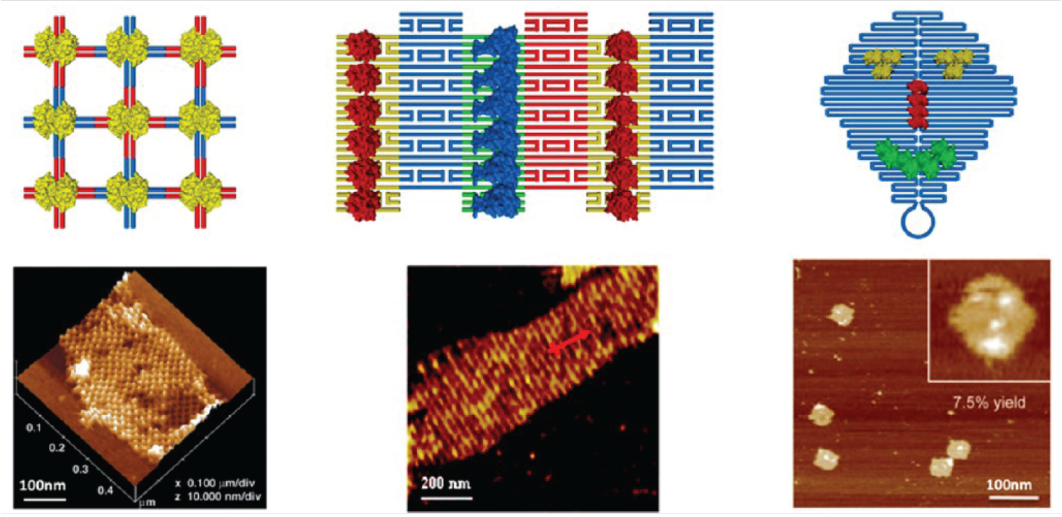

Figure 3.12: ***Representative examples of DNA-organised proteins.*** ***(Top)*** *Schematics depicting the DNA organisation of proteins, shown alongside corresponding AFM images. Adapted from Zhang et al.*[60]

small interfering RNA (siRNA)[134,135] and targeted immuno-stimulation;[136,137,138], in addition to a multitude of nanocontainers with controllable opening,[68,67,139,140] to name but a few.

In all these applications, DNA structures are utilised as templates on which additional functionality is arranged. This is easy to orchestrate for repetitive and homogeneous arrangements. An example of this can be seen in figure 3.12, where junction binding proteins are consistently attached throughout a structural DNA lattice. However, arrays of this nature offer limited functionality due to their uniformity.

In contrast, true functionality is derived from the heterogeneous arrangement of components which infer spatial asymmetry, polarity or directionality. This can be orchestrated directly through the inclusion of uniquely functionalised ssDNA staples or indirectly through the introduction of different DNA binding proteins which themselves are chemically functionalised.

These methods are, however, relatively inflexible where the number of uniquely addressable binding sites would be typically very low for any given sequence. Further, any change in desired functionality would require the re-design of the entire tile assembly, making these approaches rather costly.

An alternative approach being developed in the Bioelectronics group, is to harness the properties of homologous recombination enzymes - specifically RecA (see chapter 1) - which are able to seek out and pair DNA sequences which share homology. Hence, a standardised DNA tile could be used which would be repeatedly addressable at any unique site at any point using RecA - given

the known sequence woven throughout the structure. As this approach negates the need to pre-design binding sites or chemical functionality into the underlying DNA structure it is inherently flexible, scalable and crucially economical.[4,5,6]

3.2 RecA Mediated Homologous Recombination

The process of homologous recombination - which exchanges identical or very similar DNA strands - is critical in maintaining genomic integrity, overcoming DNA lesions, rescuing DNA replication failures and affecting genetic diversity during meiosis.[141,7,142] Consequently, it allows genetic recovery from detrimental pathways that could otherwise elicit deleterious cell activity and induce carcinogenesis.[143]

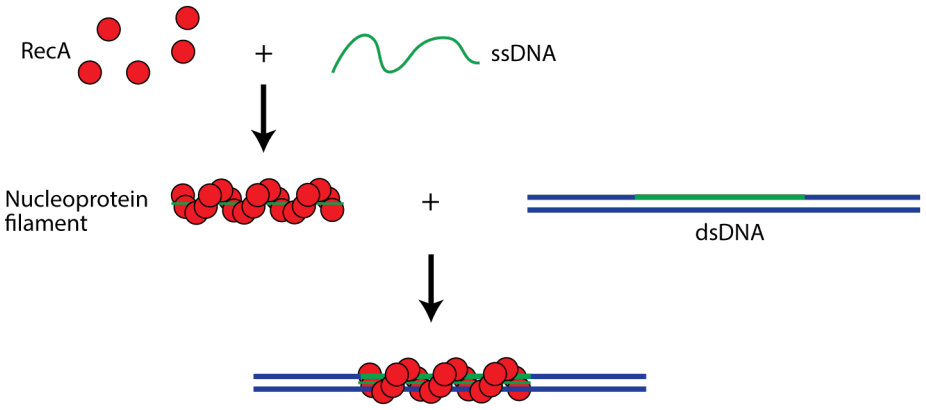

Figure 3.13: ***A schematic diagram depicting an overview of the RecA mediated homologous recombination system.*** *RecA monomers polymerised upon ssDNA forming a nucleoprotein filament. This filament is able to seek out a region of sequence homology between the encapsulated ssDNA and a dsDNA template, forming a triple stranded complex.*

Central to this process is an ATP-mediated DNA strand exchange orchestrated by ubiquitous DNA recombinase proteins - RadA in archaea, Rad51 in eukaryotes and the most widely studied homologue - RecA in bacteria.[144,145,146] As previously mentioned, RecA is the focus of this work, however, the discussion applies across all of its homologues. Furthermore, it is worth noting that there are further proteins which act at the beginning and end to activate and terminate the recombination pathway, respectively.

However within the context of this thesis only RecA is examined as it is capable of orchestrating homologous recombination in isolation. Central to this process is RecA's ability to mediate the

formation of triple stranded DNA intermediates between portions of ssDNA and dsDNA that share sequence homology, which is what makes this process of interest for bionanotechnology application. This process is shown schematically in figure 3.13.

Broadly speaking, RecA polymerises upon ssDNA in the presence of $Mg^{2+}$and a nucleoside cofactor - typically ATP - to form a nucleoprotein filament (NPF). This NPF subsequently probes a dsDNA molecule - genomic DNA in the context of a cell - for regions of sequence homology, forming a triple stranded DNA intermediate where homology is located (figure 3.13). At this point a strand exchange takes place, replacing one strand of the dsDNA with the encapsulated homologous ssDNA – creating a new heteroduplex. The exchange is finalised with the NPF complex dissociating as ATP is hydrolysed. This process will be described in more detail in the following sections.

To date, RecA and its homologues are known to have two additional functions in a cellular environment;[147,141] the regulation and induction of the SOS DNA repair pathway and the recombination of undamaged DNA during SOS mutagenesis. Although of critical importance in living cells, neither of these functions are useful for bionanotechnological application and are therefore considered beyond the scope of this thesis.

The following sections explore the structure and related functions of RecA in detail. This information is crucial when looking to harness the homologous pairing functions of RecA for spatially addressing DNA structures. Moreover, previous non-native applications of RecA are presented in justification of the methodology under development in the Bioelectronics group. Finally, a detailed examination of the approaches utilised to date to derive the current understanding of the RecA mediated homologous recombination mechanism is given, setting the scene for the investigations carried out in this thesis.

3.2.1 Recombinase A Structure

RecA is formed from the complex folding of 352 amino acids with a total calculated molecular weight of 37.842 kDa and an isoelectric point of between 5.0 - 5.6.[148] The tertiary structure of RecA consists of three distinct domains, which in turn lend themselves directly to specific functions; the large central - involved in DNA binding, the C (carboxy) and the N (amino) terminal domains.[149,18] The latter domain, is known to be of critical importance in monomer-monomer association during the filament polymerisation.[18] In contrast, the strongly negative C-terminal domain is shown to play a crucial role in modulating both duplex DNA binding and inter-filament

interactions.[150] A detailed discussion of each domain is given below. Figure 3.14 depicts the structure of the RecA protein generated from PDB entry "2REB", as reported by Story et al,[18] with the domains and secondary structure clearly highlighted.

The **large central domain** extends from residue 34 - 269 which are woven into 8 α-helices surrounded by 8 β-sheets.[18,151] Within this, there are two distinct but disordered binding loops presented on the inner surface; L-1 and L-2 which have traditionally assumed to bind ssDNA and dsDNA, respectively.[18] Due to the distinction of two binding domains, RecA is able to accommodate ssDNA and dsDNA simultaneously, which explains how it is able to mediate the formation of a triple DNA strand intermediate - central to the homologous recombination mechanism.[152] The two binding loops are shown in figure 3.14.

Due to the non-ordered and irregular nature of these loops they present a low electron density within the crystal and are therefore difficult to reconstruct. Therefore, their position is given as dashed lines connecting the flanking alpha helices in the solved crystal structure (figure 3.14). The relation of these reported structures to their functions has been confirmed by mutation studies, such as those conducted by Nastri et al[153] and Hortnagel et al.[154]

Interestingly, it is found that these putative binding loops - in addition to the flanking α-helix G of the central domain - are the most highly conserved regions of the RecA monomer across 16 different bacterial species, with 10 of the 23 disordered loop amino acids shown to be invariant across these species.[155] This inevitably supports the conclusion that it is these regions which bind DNA within the RecA monomer, a process critical to its function.

Indeed, this has been further verified experimentally through photocross-linking of the DNA species to their respective binding sites[156] and that binding can be inhibited by competition with a 20 amino acid peptide consisting of the L2 sequence.[157] However the most concrete evidence comes from point mutation studies conducted across both binding loops, which has led to the identification of the most important DNA binding residues. These are shown to be; Glycine$^{(160)}$, Glycine$^{(157)}$ and Arginine$^{(169)}$ in L1 and Glycine$^{(204)}$, Glutamate$^{(207)}$ and Glycine$^{(211)}$ in L2.[155]

Interestingly these disordered regions are found to transition to a pseudo α-helical structure upon DNA binding.[158] Moreover, the extent of α-helical structure formation is directly correlated with the presence of homologous DNA pairing compared to that of heterologous pairing. Thus the induced transition of disordered to structured binding loops maybe the mechanism by which RecA decides to transition from a synaptic to a post-synaptic complex and undergo strand exchange,[158] the process of which is presented in section 3.2.3. In addition, there is some evidence to suggest

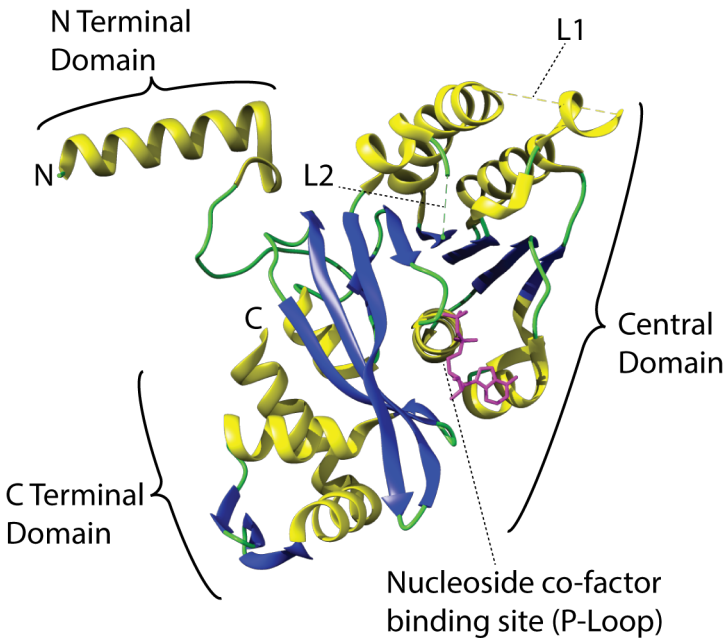

Figure 3.14: ***Crystal structure of the RecA monomer.*** *The position of the N terminal, C terminal and central domains are labelled. The binding loops for ssDNA (L2), dsDNA (L1) and the nucleoside co-factor (P-loop) are highlighted Structure rendered using Chimera, PDB file: 2REB.*[18]

that there maybe some sequence dependence, with AT rich regions inducing a greater degree of secondary structural transition.[158] This may in fact help to stabilise the homologous pairing of the weaker AT base pair bonding, bringing them in line with the stronger bonded GC regions.

More recent work by Chen et al has demonstrated the crystal structure of 6 RecA monomers bound to a ssDNA molecule. This was achieved by the fusion of the RecA monomers together – separated by flexible linkers – enabling them to form a stable nucleoprotein complex which could be successfully crystallised.[22] This work revealed that both loops L1 and L2 contribute to the binding of ssDNA and hence are not independent binding sites as previously suggested. Here, the L1 loop associates primarily with the backbone of the ssDNA by forming H bonds with the backbone phosphate groups. L2 residues from not only the binding monomer, but the two neighbours, pack alongside the nucleobases preventing their rotation away from the helical axis of the filament. In this way the ssDNA is presented with its nucleobases exposed orthogonal to and spiralling around the central axis of the nucleoprotein filament. Moreover, although it is known that RecA under-winds DNA, Chen's work has demonstrated that this occurs by stretching the DNA backbone between RecA monomers. As a result the three nucleobases held by each RecA

monomer exist largely in a B form confirmation.[22]

There are further regions within the RecA structure that are implicated in transient DNA binding, likely involved with the probing of sequence homology between the incoming and initiating strands. In particular, a tract of strongly electronegative residues which forms a cleft at the C-terminal end of the L2 loop. This region has been shown to accommodate DNA through interaction with a large proportion of basic residues.[22] These residues are also found to be well conserved throughout many bacterial species.[155] As a consequence, this region is commonly referenced as a secondary DNA binding site (site II).[159,160] Furthermore, in line with site II, DNA is shown to interact with multiple Lysine residues located on the C-terminus through the formation of salt bridges with the phosphates of the DNA backbone and hence are implicated in stabilising the initial dsDNA interaction.[161,160]

The central domain also contains a phosphate binding loop (P-loop) where residues Lysine^{72} and Threonine^{73} are found to be directly responsible for the hydrolysis of ATP and its homologues,[162,163,151] interacting directly with the nucleosides phosphate group. Under physiological conditions, RecA has a typical ATP turnover rate of 30 min^{-1} and 20 min^{-1} on ssDNA and dsDNA, respectively.[164] Indeed, it has been shown that RecA mutants that have either of these residues substituted continue to be able to bind, but lack the ability to hydrolyse ATP to ADP.[165,166,167] Taken together, these findings describe RecA as being a DNA dependent ATPase and indicate how the complex is locked in the presence of ATPγS - a non hydrolysable analogue.

The **N - Terminal domain** consists of residues 1 to 33 which form a protruding large α-helix and short β-strand.[18] This largely negative α/β unit is crucial in the monomer-monomer associations that enable polymerisation of RecA to occur.[18] This is examined in detail in section 3.2.2.

In contrast, the **C - terminal domain** is considered to begin from residue 270 until 352, consisting of 3 α helices and 3 short β sheets. This domain, is suggested to define the active or inactive state of the polymerised complex, in two ways.[168] Firstly, the state of the RecA complex is defined by the bound nucleoside cofactor; ATP (or ATPγS) for the active or ADP for the inactive. It was noted by Yu et al[168], that the position of the C-terminal domain relative to the central domain altered depending on the state of the complex. Interestingly, the C-terminus is noted for its high concentration of negatively charged residues which are capable of repelling negatively charged dsDNA molecules.[169,150] When taken together, these findings suggest that the C-terminal domain acts as a gateway to the binding domains within the RecA polymer, blocking the entry of an invading dsDNA molecule when the complex is in an inactive state.

Furthermore, the C-terminal domain is implicated in formation of inter-filament interactions observed between neighbouring RecA polymers, binding close to the N-terminal domain of neighbouring RecA polymers.[18,151] These filament bundles are observed *in vivo* and may act as a storage mechanism for inactive RecA, inhibiting its ability to bind spuriously to dsDNA by blocking entry into the helical groove of the protein complex.[18,150,170] Mutation studies have implicated, Glutamate$^{(38)}$ and Isoleucine$^{(298)}$ to be critical to this process, where substitutions inhibit the ability of RecA to form bundles, whilst promoting increased dsDNA binding thus supporting the role of the C-terminal domain in both the above functions.[170,171]

3.2.2 RecA Nucleoprotein Filament Structure

Central to the RecA mediated homologous recombination mechanism is the nucleoprotein filament formed from the polymerisation of RecA monomers upon DNA. A section of a RecA NPF can be seen in figure 3.15 A. It is also important to note, that RecA is capable of forming "inactive" polymers in the absence of DNA. These are generally accepted to have a compressed geometry compared to that of the "active" (DNA containing) form. The formation and action of this polymer is discussed below, prior to which the structure is described.

In the "active" state - where ATP or ATPγS is included - RecA is bound to DNA occupying 3 nucleotides per monomer. The resultant nucleoprotein filament is 120 Å wide with a 2.5 Å central cavity, along which the incorporated nucleobases are exposed (figure 3.15 A). The NPF wraps around in a helical conformation with 6 RecA monomers per turn, giving rise to a right-handed helical pitch of 95 Å with an axial rise of 5 Å (figure 3.15 A).[20,19,21] As a result, the nucleic acid is held along the central axis with a relaxed helical pitch of 18.5 nucleotides per turn, approximately 1.5 times longer than native B-form dsDNA.[172,173,159] This elongation leads to the notion that the RecA acts as a structural scaffold holding the DNA in an extended conformation for the strand exchange to take place.

This observation is further supported by the formation of free RecA hexamers which display the same geometry when imaged with electron microscopy.[174] This hexameric secondary structure occurs in the absence of DNA, thus it can be considered that within the nucleoprotein filament it is the RecA that supports the observed structure rather than the DNA.

In contrast, the "inactive" state is characterised by a compressed helical pitch of 73 - 83 Å, with otherwise identical hexameric symmetry.[18] Despite a 10 Å discrepancy between the observations of Story et al[151] and Egelman et al,[21] their reported structures are generally accepted to represent

the same RecA polymer inactive state, with differences attributed to the method of structural derivation, X-ray crystallography in the former and electron microscopic reconstruction in the latter.

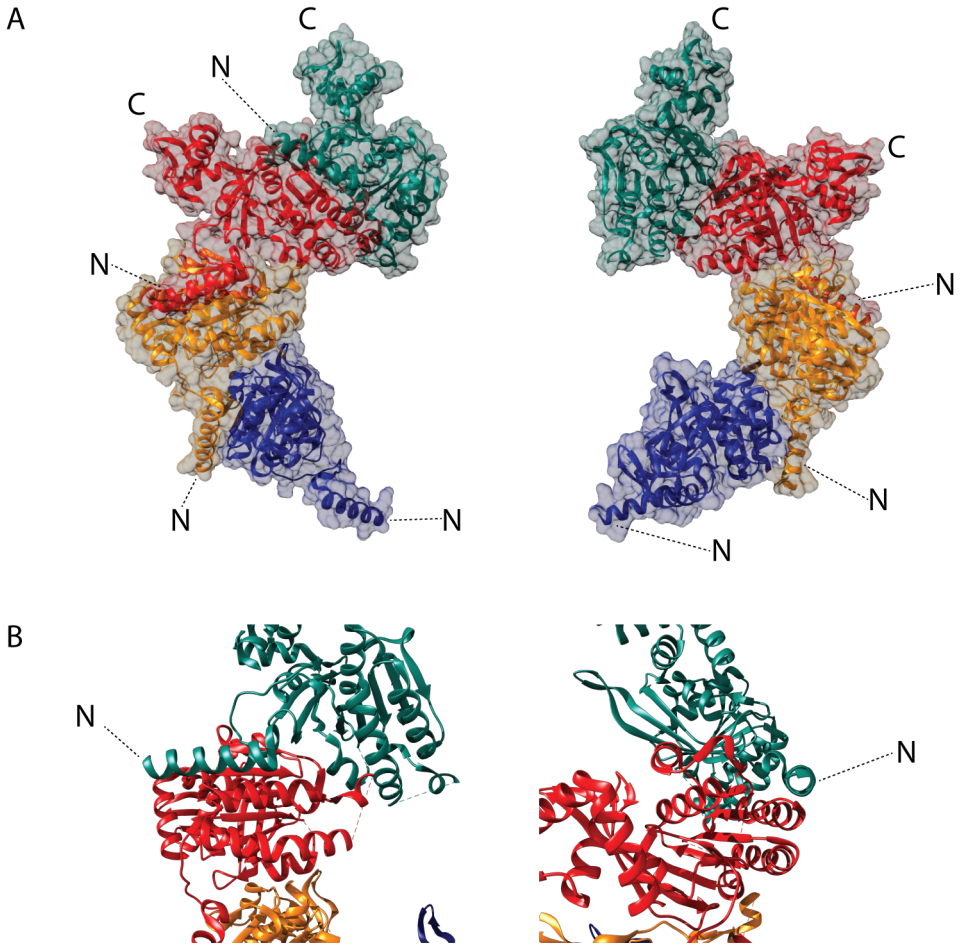

Figure 3.15: ***The crystal structure of a RecA nucleoprotein filament.*** ***(A)*** *The individual monomers in a nucleoprotein polymer are highlighted in different colours for clarity. The structure is rotated by 180°, left to right. The solvent accessible surface is shown as a semi-transparent layer, beneath which the protein structure is represented in a ribbon model. The respective locations of the N and C terminal domains are shown for each monomer.* ***(B)*** *A close up view depicting the association of the N terminal domain of one monomer with the large central domain of another, forming an a/b (parallel) unit. Shown from the top (left) and rotated 90° (right). Structure assembled in Chimera, from PDB file: 2REB.*[18]

In these structures, the C-terminal domains are found to project out from the right-handed helical arrangement in a radial formation (figure 3.15 A); thus suggesting the proposed involvement in inter-filament interactions and mediating DNA binding (section 3.2.1).[150,170,171] In contrast, the N-terminal domain is found to be critical in monomer-monomer associations, where the most critical residues are Lysine[216], Phenylalanine[217] and Arginine[222], as identified by mutation studies.[175] In

this process, the N-terminal α helix of one monomer packs in between the α helix and β sheet of the adjacent monomer's central domain (figure 3.15 B), causing a reduction in the solvent accessible surface area by 2.9 nm^2. This creates a stable dimer interface, held together by hydrophobic and electrostatic interactions. Furthermore, the importance of this interaction is evidenced through the inhibition of polymer formation in RecA mutants that have had their N-terminal domains completely truncated.[175]

3.2.3 RecA Mediated Homologous Recombination

Having examined the structure of the RecA monomer and NPF it is important to discuss how that structure relates to the mechanism of homologous recombination.

Broadly speaking, the process of homologous recombination as mediated by RecA can be considered to consist of three distinct stages;[176,177]

- A Pre-synaptic phase - the formation of a nucleoprotein filament
- A Synaptic phase - the location of homology
- A Post-synaptic phase - the exchange of DNA strands

The pre-synaptic phase: RecA in the presence of ATP (or its analogues) and Mg^{2+}binds to a ssDNA or dsDNA - depending on the cofactor concentration - at a stoichiometry of 3 nucleotides (nt) per monomer (see section 3.2.2) forming an active nucleoprotein filament, also known as the pre-synaptic filament.[178] The filament formation process occurs in two phases; a rate limiting nucleation step and a extension step (figure 3.16).[7]

Nucleation: Firstly, free RecA monomers nucleate upon the DNA, typically taking 6 - 8 monomers to initiate full polymerisation.[7] This step is found to have a pH dependence and is critically enhanced in the presence of divalent ions, with a particular prevalence for Mg^{2+}.[7,179] Interestingly the preference for polymerisation on ssDNA - RecAs primary target - and dsDNA is governed by the concentration of Mg^{2+}, low and high, respectively.[7]

RecA nucleation has been shown to be severely hindered by the presence of secondary structures within the target DNA.[180] Typically this limitation is circumnavigated by accessory proteins such as single stranded binding protein (SSB) which removes secondary structure formation in ssDNA and prevents its premature annealing back into dsDNA.[181] Hence, SSB critically reduces the energy barrier for RecA binding, leading to high nucleation rates.[179] Furthermore, SSB may

also be responsible for ensuring the formation of continuously and regularly polymerised RecA filament structure by pre-straightening the DNA template. This enhanced RecA loading rate at regions of disrupted dsDNA is further noted in its apparent preference for unusual DNA topologies such as Z form DNA.[182]

Extension: The second stage occurs with a rapid extension in the 5' to 3' direction, with monomers binding at the growing 3' polymer interface and dissociating at a slower rate from the 5' end, with an approximate K_D = 100 nM.[182,183] Interestingly, where dsDNA is concerned, binding occurs to one of the strands in the duplex. Hence filament growth can occur from either 5' termini of the duplex DNA and extend accordingly.[164] It is important to note that the polymerisation of RecA is not sequence dependent. However, preferences for GC-rich regions of DNA have been demonstrated where increased rates of extension are notable.[184]

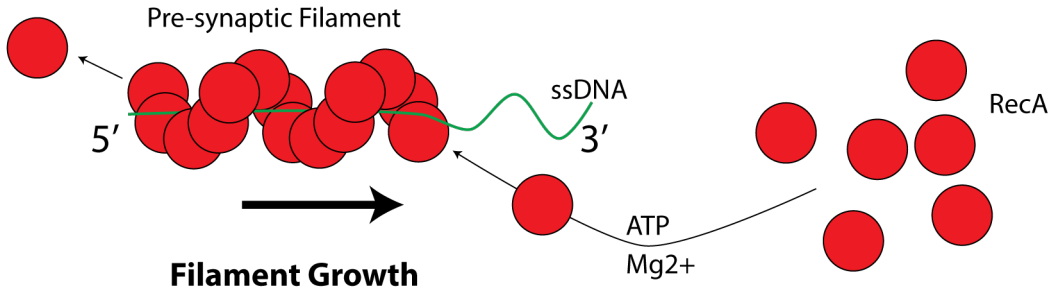

Figure 3.16: ***A cartoon depicting the formation of a RecA nucleoprotein filament.*** *RecA monomers nucleate upon ssDNA in the presence of co-factors. Monomers polymerise at the 3ı face and disassemble from the 5ı face at a slower rate. Net filament growth occurs in a 5ı to 3ı direction.*

As mentioned, filament growth is a process of net gain, where RecA monomers are arriving at the 3' termini and leaving from the 5' simultaneously (figure 3.16).[185,186] The rate of disassembly is linked directly to the to ATP hydrolysis, with a disassembly rate of approximately 70 or 120 monomers min^{-1} on ssDNA and dsDNA, respectively at $37°C$.[186] Note that disassembly is largely suppressed where ATPγS (see section 1) is utilised and hence the filament growth rate is accelerated.

The resultant pre-synaptic filament is central to the RecA protein function containing an initiating ssDNA which is able to incorporate an invading dsDNA as it searches for homology and ultimately catalyses a homologous pairing between the two DNA molecules. Interestingly in this way, it is considered that the nucleoprotein filament acts to spatially confine the interacting DNA species, enabling strand exchange to take place, rather than enacting it directly.[177] Once formed, the entire nucleoprotein filament complex undertakes a homology search, probing the sequence of

a dsDNA molecule in a synaptic phase (figure 3.17). Note that a distinction must be made between homology probing where base pairing is tested and homology searching which is the larger movement of the filament in finding homology (figure 3.17). The former is described below and the latter remains a widely debated topic, where a combination of random sampling, inter-segmental transfer - dependant on NPF size - and facilitated 1D diffusion along the DNA is argued (figure 3.17 B).

Synaptic phase: In this phase, the pre-synaptic filament aligns with a given dsDNA and proceeds to probe for regions of sequence homology with the encapsulated initiating ssDNA .[187] It is critical to note that this process occurs in the absence ATP hydrolysis and hence no known motor function of RecA has been identified.

Evidence from molecular dynamics (MD) simulations suggest that initial contact occurs through weak interactions at the protruding RecA C-termini (section 3.4.2).[161] From here the dsDNA diffuses into the filaments helical groove, docking into binding site II through the formation of salt bridges to the phosphate backbone. In this position, the dsDNA is destabilised due to the induced tension as it forms a kink.[188,160] This allows residues from the L2 disordered loop to intercalate with the dsDNA, further destabilising the local base stacking causing the characteristic relaxation and extension of the dsDNA from B-form.[160]

At this stage both DNA molecules are bound simultaneously in an under-wound configuration within the RecA filament complex.[12,189] Here the complementary strand is now positioned to undergo base flipping to probe for base pair homology with the initiating ssDNA strand.[190,11] Where homology is located, stable base flips result in the loss of backbone salt bridges from site II and a re-arrangement of the backbone into the L2 position stabilising the new pairing.[160]

Such base flipping was traditionally assumed to occur with a minimal unit cell of 3 bases as this is the number of nucleotides bound per RecA monomer.[191,16] As a result, this process is purported to be slow as it requires the movement of residues in the L2 loop to allow full triplet flipping, thus avoiding steric hindrance.[161] Despite this interpretation, this doesn't quite fit with the observed kinetics, which are found to scale non-linearly with the size of the nucleoprotein filament.[192,193] Indeed, previous work has demonstrated that 6 - 8 nt homology is sufficient to form a post-synaptic complex.[6,194,195]

In contrast, recent work by Yang et al suggests that base flips occur in duplets, running in the 5' to 3' direction.[160] Here, the third base does not flip due to the aforementioned steric hindrance and is likely to flip into place consolidating the final sequence upon dissociation of the RecA complex.

Interestingly, new stable base pairings are only noted to occur in pairs, due to the effect of stacking interactions and thus single spurious base flips do not result in new watson-crick interactions.[160]

Moreover, base flipping is suggested to occur surrounded by the hydrophobic residues of the L2 loop as a consequence of their destabilising intercalation. Thus flipping proceeds continuously protected from the environmental water within a favourable hydrophobic environment. As a consequence the energy barrier to base flipping is drastically reduced compared to a base flip that were to occur through thermal fluctuations alone, having to pass through an unfavourable aqueous environment from one stable configuration to another.[160] These observations go some way to explaining the apparent kinetic paradox of a genomic wide homology search, suggesting homology probing requires significant less energy than previously discussed.

Homology probing is believed to continue to processes in this fashion, forming an increasingly stable new heteroduplex where homology has been located.[196] Upon *de novo* pairing of the entire sequence homology, the complex is considered to have transitioned to a post-synaptic phase.

The Post-synaptic phase: On location of a homologous region, a strand exchange between the two DNA molecules ensues where the complementary strand of the incoming dsDNA molecule is moved entirely into the L2 position, forming a new stable heteroduplex with the initiating ssDNA strand (figure 3.17).[197,198] At this stage, ATP hydrolysis occurs causing a conformational change in the RecA forcing out the remaining incoming DNA strand and disassembling the nucleoprotein filament.[191,199,200] This stage of the homologous recombination mechanism can be stalled through the use of non-hydrolysable analogues of ATP, such as ATPγS due to the severe reduction in the nucleoside turnover rate and hence a stable triple stranded complex can be formed.[201] It is important to note, that during this final stage further supporting proteins are known to be involved in order to resolve the complex fully, such as the operation of DNA polymerase to repair the missing region from the donor dsDNA .

Work to date, delivers a reasonable description of the RecA mediated homologous recombination mechanism, including; the formation of the nucleoprotein complex, the homology probing via the flipping of bases between the incoming and initiating DNA strands, and the disassembly of the complex, completing the formation of the new heteroduplex.

However, many questions remain regarding the process by which the RecA filament is able to search through large regions of dsDNA, for example in the context of genome wide search.[202,187] It is important to remember that this “homology searching” process occurs in the absence of any motor activity form the RecA itself. Furthermore, although it is known that a large collection of

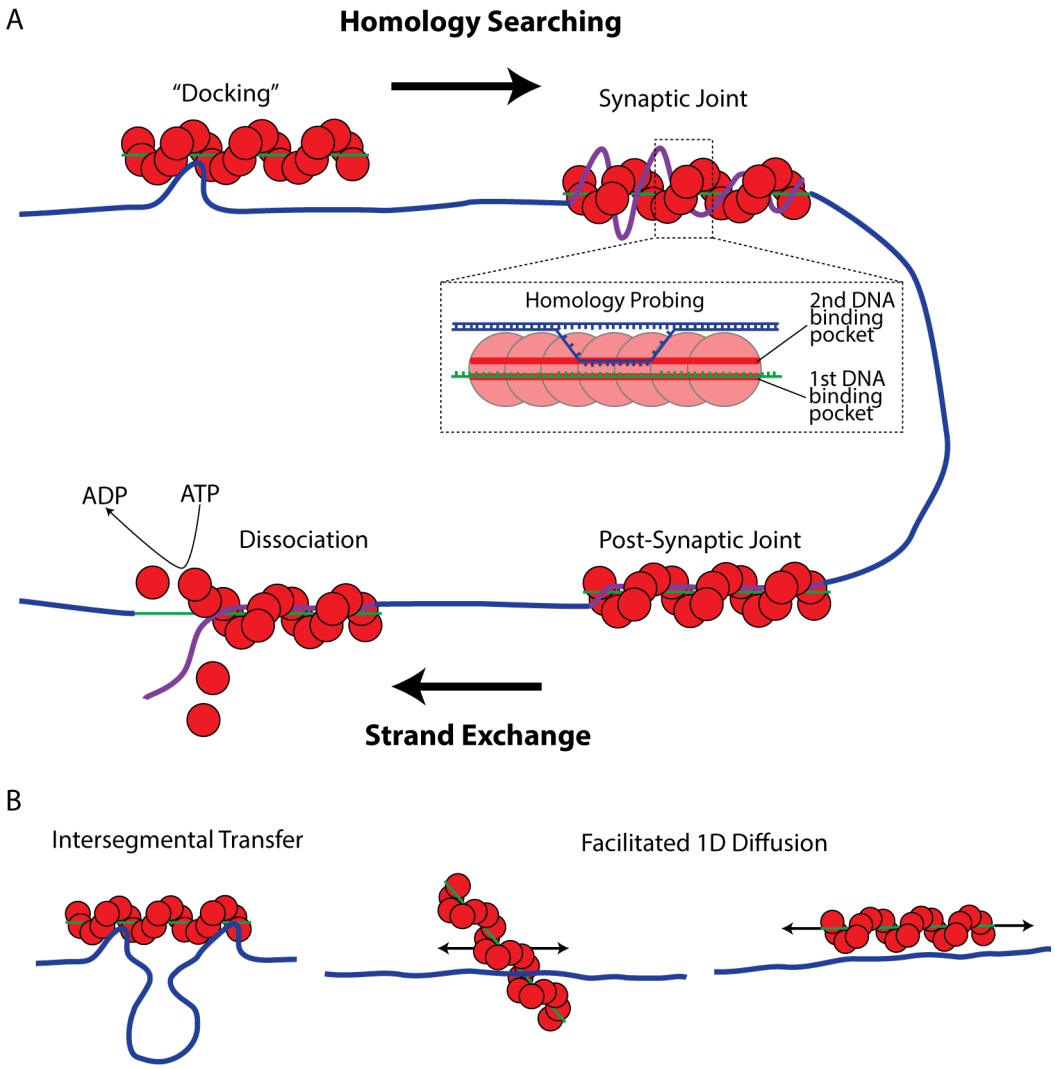

Figure 3.17: ***A cartoon depicting a detailed overview of RecA mediated homology searching and strand exchange.*** ***A*** *Nucleoprotein filaments "dock" with dsDNA through random contact. Homology search is undertaken, forming a pre-synaptic joint as sequence homology is probed as the dsDNA is transiently bound in the secondary binding pocket. Upon location of homology, the incoming dsDNA is bound to the L1 binding loop, forming a stable Post-synaptic complex. Strand exchange occurs as ATP is hydrolysed and the RecA monomers dissociate resulting in a the formation of a heterogeneous dsDNA strand.* ***(B)*** *Homology searching through transient dsDNA association is postulated to occur through, Inter-segmental transfer between distal non-contiguous regions and facilitated diffusion along the template molecule.*

protein act in support of the process, RecA is demonstrated to be capable of enacting homologous recombination in isolation. Hence, it must be concluded that no motor activity or direction is derived from the support proteins either. Thus the mechanism by which RecA nucleoprotein filaments are able to search genomes for regions of homologous sequences remains an area of

intense study.[177,203,204,202]

3.3 DNA:RecA Nucleoprotein Complexes as a Construction Material

The sequence specification and high fidelity of the RecA mediated DNA triple-strand formation are attractive characteristics when considering self-assembling nano-structures for bionanotechnology applications. Furthermore, it has been shown that the functionalisation of RecA allows for specific material arrangements to be made using DNA as an underlying scaffold.[9] Such methodologies commonly refer to the formation of site specific triple-stranded DNA junctions by RecA as RecA mediated patterning, and shall be referred to henceforth, where relevant, in this text. In addition to this, there have been multiple demonstrations that have harnessed the sequence specificity of RecA to enhance the precision of traditional molecular biology techniques. However, as this thesis is concerned only with bionanotechnological application, these are not reviewed here. The following section aims to review the extant literature on the technological applications of RecA.

3.3.1 Bionanotechnology Application: Molecular Lithography and Site Specific Metalisation

It has previously been demonstrated by Braun et al,[205] that continuous nanowires can be formed by nucleating Silver ions onto glutraldehyde functionalised DNA. Here, Silver nanoparticles are bound to the backbone of λ bacteriophage DNA (48,502 bp) at alkaline pH resulting in a continuous wire 30 - 50 nm in diameter. Successive work improved the conductivity of the wires through the sequential nucleation of Gold on the Silver nanowires, following the well established photographic process of Tollen's chemistry.[10]

In addition, incorporation of RecA mediated molecular lithography was demonstrated, representing the first practical application of RecA mediated DNA assembly for bionanotechnology. Here, a 2027 nt RecA nucleoprotein filament was reacted with bacteriophage λ DNA that had been pretreated with glutraldehyde. This 500 nm region of DNA was subsequently shielded from exposure to the successive Silver and Gold nucleations by the RecA and hence a break was specifically created in the nanowire upon its removal (figure 3.18).[10]

Further work by Keren et al[9] reported the arrangement of functionalised carbon nanotubes (CNTs) through antibody attachment to RecA nucleoprotein filaments, to produce a field effect transistor.

Here, the DNA was pretreated as previously and a RecA nucleoprotein filament site specific arrangement formed. This positioned nucleoprotein filament was targeted with biotinalated anti-RecA antibodies, and the whole construct subsequently exposed to Streptavadin functionalised single walled carbon nanotubes (SWNT). This construction enabled the specific placement of a single carbon nanotube aligned along the length of the DNA template. Finally, the exposed DNA was metallised with Silver and Gold, as previously, resulting in continuous gold nanowires directly contacting either termini of the single SWNT (figure 3.18).[9]

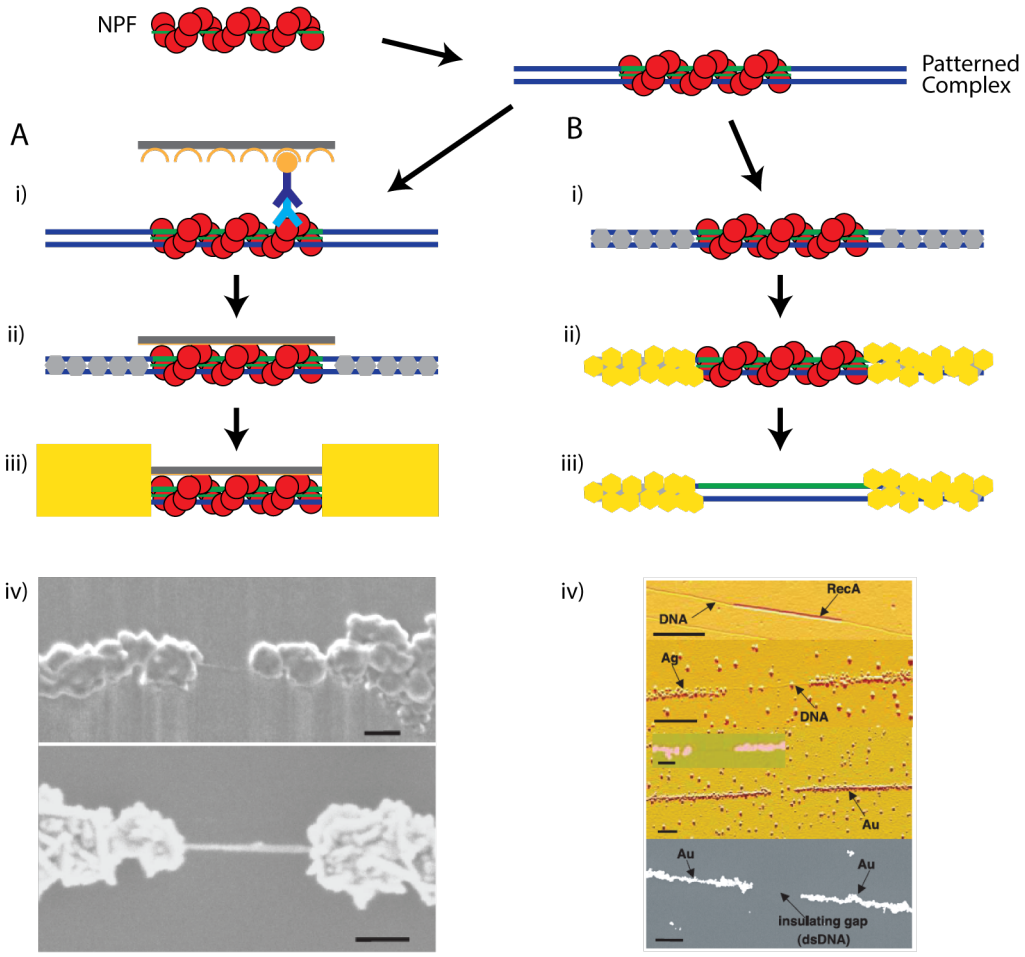

Figure 3.18: ***RecA mediated molecular lithography and templating.*** *Nucleoprotein filaments and interacted with dsDNA . **(A)** formation of a field effect transistor by specific placement of SWNTs. **(i)** A functionalised SWNT is attached to the RecA nucleoprotein filament through a conjugated antibody mechanism. The surrounding DNA is selectively metallised, using Ag **(ii)** and Au **(iii)**. The RecA locally protects the dsDNA from metallisation, leaving the SWNT bridging an otherwise insulating gap. **(iv)** SEM images depicting the final FET. **(B) (i) & (ii)** RecA locally protects the dsDNA from metallisation forming nanowires. Removal of the RecA results in specifically positioned insulating gaps **(iii)**. **(iv)** AFM and SEM (bottom) images depicting the successful formation at various stages. Adapted from Keren et al.*[10,9]

Other RecA mediated site-specific arrangements of gold nanoparticles (GNPs) have been demonstrated by Nishinaka et al.[8] In this work DNA templates were nucleated with a RecA mutant containing an additional cysteine residue at the radially protruding C-terminus.

Subsequent introduction of gold nanoparticles (GNPs) - which bind selectively to the exposed cysteine residues - culminated in the successful fabrication of nanowires exclusively along the RecA nucleoprotein filament. This methodology was used to successfully fabricate alternating conducting and insulating regions upon DNA templates through the successive binding of wildtype and cysteine modified RecA nucleoprotein filaments onto neighbouring regions. In addition, the same group demonstrated the formation of conducting junctions, through the interaction of partially homologous nucleoprotein filaments. This resulted in only short portion of the RecA filament forming a triple stranded region with the template DNA. Addition of further RecA, coating the remaining DNA, and subsequent exposure to GNPs resulted in the formation of a fully metallised three arm junction.[8]

The work discussed above, successfully demonstrated that self assembled RecA structures could be a legitimate alternative to top down lithographic fabrication for electronic structures. However, despite this impressive work, the large feature sizes produced make them uncompetitive with current photolithographic techniques. Furthermore, to date, work has focused upon 1D geometries limiting the applicability of such systems. Thus further advancements require substantial growth in 2 and 3 dimensions for these techniques to be adopted as a plausible alternative for the arrangement of electronic circuits and architectures.

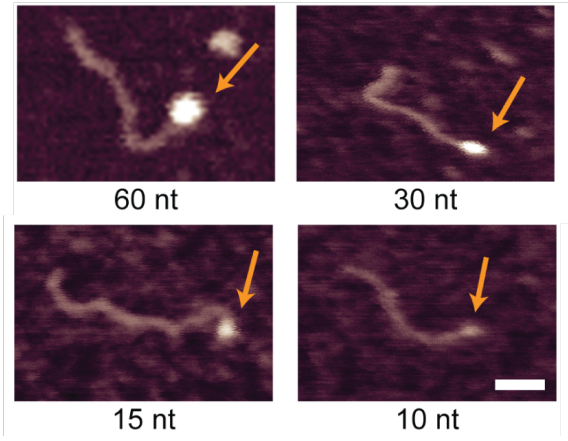

Figure 3.19: ***Scaling down RecA mediating patterning.*** *A series of AFM images showing RecA nucleoprotein filaments of decreasing size, formed upon single-stranded DNA from 60 nucleotides (60 nt) to 10 nucleotides (10 nt) long, patterned at the termini of a double-stranded DNA template (orange arrow). The scale bar indicates 50 nm. Adapted from Sharma et al.*[6]

In response to this, work by Sharma et al examined the extent to which RecA mediated patterning could be efficiently scaled down (figure 3.19) and conducted in parallel, that is to say if multiple regions can be patterned simultaneously on the same DNA substrates.[4,5,6]

Here, RecA mediated patterning was successfully demonstrated at the termini of DNA substrates with nucleoprotein filaments down to 6 nt in length, representing a feature size of 3 nm. In contrast a lower limit of 20 nt was found for nucleoprotein filaments patterned at the center of a DNA substrate. Nonetheless, this still represents a programmable feature size of 10 nm - remaining competitive with the most advanced photolithographic processes to date.[6] Furthermore, simultaneous and co-operative patterning upon the same DNA substrate at the scales presented above has also been reported. This work demonstrates that at least 3 nucleoprotein filaments can be simultaneously patterned independently of one and other.[5]

As a consequence of the work by Sharma et al,[4,5,6] the work of Keren et al[10,9] and Nishinaka et al[8], this RecA mediated approach has been highlighted as a potential competitor to top-down photolithographic techniques. However, further work is required to adapt these approaches from demonstrations of linear DNA substrates to more complex topologies, such as those demonstrated in DNA nanotechnology (section 3.1).

In order to achieve this, a deeper understanding of how RecA mediates the location of homologous DNA sequences must be obtained, such that complex DNA structures maybe designed with out undermining the achievable patterning efficiencies of the system. Visualising the resulting small structures is of great importance to understand and exploit this process, however, given the typical feature sizes and the structural properties of the materials involved, it remains a significant challenge. The next section will examine the interrogation approaches that have led to our current understanding of the RecA mediated homologous recombination mechanism and how these might be improved upon to gain further insight.

3.4 On the Interrogation of the RecA Nucleoprotein and its Dynamics

The apparent discrepancy in our understanding regarding the RecA mediated homology search maybe considered a direct consequence of the employed investigation methodologies, the majority of which have employed ensemble methods. For example; our current understanding of the RecA structure, both monomeric[18] and polymeric comes from x-ray crystallographic techniques, small angle neutron diffraction (SAND) and electron tomography (ET).[206,207,208,22,19] Using these averaged structures, attempts have been made to decode the interactions of RecA with DNA (both single stranded and duplex) in an effort to identify intermediates that can be used to characterise

the mechanism. Such work, as that conducted by Egelman et al (figure 3.20 A)[19,20,21], Datta et al[149], Story et al[18] and Chen et al[22], give indispensable structural insight when coupled with the correlating biochemical data in the literature[13] and have provided a starting point for molecular dynamics simulations. However due to the differential fixation and sample preparations for all these techniques, discrepancies and scepticisms arise as to whether these are true depictions of nucleoprotein structures and if the structural changes identified can actually be attributed to stages within the interaction mechanism.[21,12]

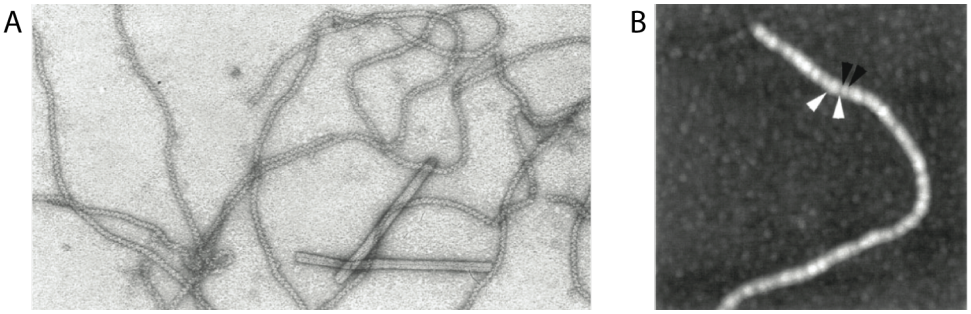

Figure 3.20: ***Microscopy examination of RecA nucleoprotein filaments.*** ***A*** *Electron microscopy image of negatively stained RecA:dsDNA:ATPγs nucleoprotein complexes.* ***B*** *AFM topographical images of RecA:dsDNA:ATPγs nucleoprotein complexes acquired with CNT tips. Filaments have been fixed with 0.1% Glutralahyde. White and Black arrows indicate 20 and 10 nm periodicities, respectively. (A) Adapted from Egelman et al.[20] (B) Adapted from Umemura et al.[209]*

In contrast, the AFM offers many advantages with respect to elucidating the structure and interactions occurring within RecA mediated nucleoprotein filaments. Most importantly, biological molecules are maintained without staining or fixation, thus preserving their natural geometries, structures and conformations that are the specific point of interest.[210] Furthermore, imaging can be conducted in air under ambient conditions or in aqueous solutions. In this case the latter is applicable to understanding the geometry of NPFs *in vivo* and the former remains true to their intended use as integral parts of a nano-circuit, hence dried upon a surface. Of particular interest is the resolution obtained by Umemura et al[209] using carbon nanotube (CNT) probes (figure 3.20). By this method, a 10 nm helical pitch for NPFs formed upon dsDNA was observed, found to be in good agreement with that predicted from crystallographic methods.[20] Furthermore they reported a sufficient aspect ratio such that a groove of 1 nm was determined to exist at the pitch between helical stacks of RecA.[209] Interestingly, from our earlier work, it was found that through rigorous purification and sample preparation that almost equivalent spatial resolution could be gained with commercially available probes.[211] Despite these studies, significant resolution enhancement is still desired on isolated RecA nucleoprotein filaments to

discern the intermolecular intricacies which are undoubtedly of key importance to understanding the function of RecA, this is examined in chapter 9.

3.4.1 Filament Formation and Disassembly

Traditionally, assessments of RecA interactions, both monomeric and as a nucleoprotein complex, have been conducted using biochemical assay[182,12,212,13], providing valuable but limited insights into these mechanisms. Initially monomeric RecA interactions with both double stranded and single stranded DNA were the focus of these previous studies, which is a trend that extended into single molecule experiments as well.[183,212,213,214,178]

Such work as that by Pugh et al[182] adapted the tau analysis, also known as the abortive initiation assay, which was originally developed for studying differential kinetic stages of RNA polymerase association with an open promoter complex. Here, it was utilised to identify two stages in the binding of RecA to dsDNA with distinctive kinetics: a rate limiting co-operative nucleation step and a rapid polymerisation step, presented in section 3.2.3. Further analysis indicated that these steps were differentially perturbed by environmental factors.

Thus, biochemical analysis has laid the ground work for assessing the biological functions of RecA and its complexes, however a more direct approach towards single molecule visualisation is necessary. Work by Bell et al[215] and Joo et al[183] attempted to address this using fluorescently labelled RecA to directly detect the nucleation and growth steps upon ss- and dsDNA, using total internal reflected fluorescence (TIRF) and Forster resonance energy transfer (FRET), respectively. It is interesting to note that although these techniques remain bulk analysis, Joo et al[183] were able resolve signal fluctuations indicative of the assembly of individual RecA monomers with their FRET set up.

Adding additional precision to these measurements, Conover et al[216] and Arata et al[217] utilised magnetic tweezers to detect the known relaxation of the DNA helix in real time as a consequence of RecA polymerisation. Interestingly, Conover et al[216], noted that application of tension or compression to the dsDNA strand severely hindered the ability of RecA to form a nucleoprotein filament. This indicates that RecA interacts with DNA preferentially when it has access to its full degrees of freedom, free in solution and that the appearance of condensates or supercoiling hinders the monomeric interaction. As will be examined later, this is a theme found to be true of the NPF:DNA interaction also. Moreover, using a new generation of magnetic tweezers that were able to detect a twisting/rotation motion of less than 5°, Arata et al[217], described individual

rotation steps of 60° due to the sequential monomeric association of the Rad51 protein. Although it is noted that the under-winding of DNA by Rad51, a human homologue of RecA, occurs to a lesser extent than that of its *E.coli* counterpart, with a NPF helical rise of 4.5 Å compared to 5.1 Å. Interestingly, where a torque was applied the polymerisation was slowed and they were able to drive the dissociation of Rad51 causing the DNA to return to its B-DNA form.

Although these studies can provide information about the mechanics and kinetics of RecA polymerisation, they are little substitute for a true microscopical characterisation of in situ events. To that end the AFM has been utilised to characterise the assembly and disassembly of RecA nucleoprotein filaments upon dsDNA *in situ*.[218,219,213,185] These studies indicate the applicability of following, in real time, the interaction of RecA and DNA, in support of static disassembly evidence observed previously in the authors masters thesis.[220]

3.4.2 Nucleoprotein Filament Interactions with dsDNA

In contrast, when examining the dynamics of NPF and DNA associations, very few direct observations have been attempted. The majority of understanding is derived from fixed or static examinations of reaction intermediates and derived complexes, as discussed above.[19,20,206,207,208,149,18] Although these observations have provided an insight and a useful starting point for molecular modelling, many of the differences in observed structures tend to suggest a degree of sample preparation artefacts.

As with the monomeric and polymerisation studies, efforts began through the use of cleverly designed biochemical assays.[12,13] Of particular interest is the study conducted by Adzuma et al[13], to identify or disprove the existence of a sliding motion occurring during the NPF search for homology. Here, a plasmid template containing 4 identical sequentially repeated targets was used. They hypothesised that if a sliding mechanism was the main method of homology search then postsynaptic joints would form predominantly on the outer two targets. However, if the interaction occurred through a random association, then there would be no preferential target binding and each region of homology would have an equal likelihood of containing a postsynaptic joint. The main conclusions drawn from this work was that sliding does not occur to any “significant extent”. However, due to the design of the experimental template, a significant extent is defined as nearly 1 kbp. Although highlighting that facilitated diffusion along a DNA template over 1000’s of base pairs is not energetically favourable, an observation that is supported by others[221], it is not however conclusive in ruling out a sliding motion as part of an overall, more complex mechanism.

Elsewhere, specific attempts to observe the interaction of nucleoprotein filaments and dsDNA in real time have been conducted successfully in bulk using fluorescently labelled NPFs.[222,204,11] One such study by Van der Heijden et al[14], utilised magnetic tweezers to study the kinetics of the NPF invasion. Here the interaction of NPF to a tethered dsDNA template caused the characteristic local relaxation at the region of homology, giving rise to a measurable increase in length and rotation of a magnetic bead. Although this methodology could not elucidate the interaction mechanism, it does begin to highlight the interaction kinetics, which scale roughly linearly with the NPF length. Here, for example, a 250 nt NPF was found to take 100 seconds to interact.

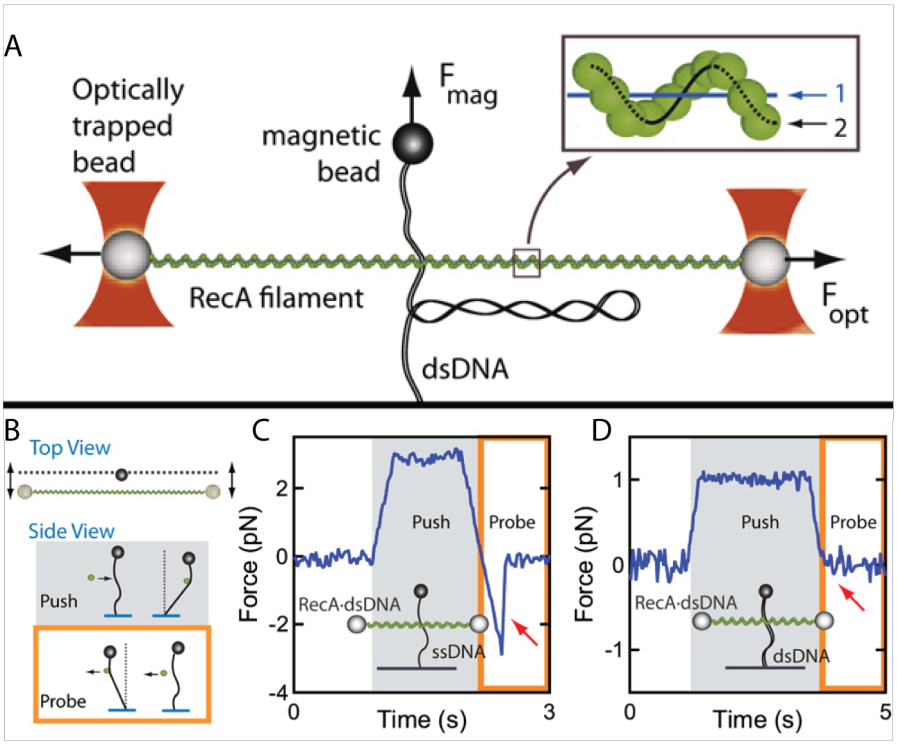

Figure 3.21: ***A dual molecule experiment for studying homology recognition during RecA homology searching.*** ***(A)*** *RecA filaments and DNA molecules are trapped in multiple orientations using optical and magnetic tweezers. Movement of one relative to the other enables detection of various interaction forces, Fmag and F opt.* ***(B)*** *Interaction forces between nucleoprotein filaments and DNA can be detected by pushing the two molecules together and measuring the restoring forces as they are separated (red arrows in traces* ***C*** *and* ***D****). Adapted from DeVlaminck et al.*[188]

Moreover, work by Forget et al[11], utilised optical tweezers to manipulate the tension in a target dsDNA to examine the impacts on RecA protein or NPF interactions. It is proposed that the interaction with a dsDNA molecule requires access to the full three dimensional degrees of freedom associated with DNA in solution. Where DNA condensation occurs, these interaction processes are hindered either due to the introduction of compression or tension to the dsDNA

helix. The work by Forget et al, supports this by allowing fluorescently labelled nucleoprotein filaments to interact with a trapped dsDNA molecule when relaxed in solution. Note that here, both ends of the dsDNA molecule were attached to different latex beads, which were manipulated using optical tweezers, allowing tension to be induced in the DNA strand. Interestingly, this acts to cap the ends of the DNA, ruling out the necessity for the nucleoprotein filament to associate with a free termini or break in the DNA target. The authors proposed that the 3D degrees of freedom of the DNA substrate are of key importance to the homology searching interaction and further suggested that the occurrence of inter-segmental transfer between distal portions of a non-contiguous DNA substrate go some-way to explaining the apparent searching efficiency within a complex genomic environment.[11] Both of these single molecule observations with optical[11] and magnetic tweezers,[14] have demonstrated second order kinetics with rate constants on the order of 100's of seconds for RecA homology searching.

More recently, FRET based techniques have been employed for studying the RecA mediated interactions between short DNA molecules.[15,16,17] In particular the work by Ragunathan et al has provided evidence for the existence of a 1D facilitated diffusion along the contour of the dsDNA molecule, demonstrated here for short filaments (80 nt) which were shown to continually search up to 300 bp regions of dsDNA before dissociating.[15] As demonstrated by Adzuma et al, this progressive motion was however ruled out for long-range slides (several kbp).[13] Nevertheless, relatively stable synaptic association must be required to form for such a concurrent slide and progressive search process.

Perhaps *in vivo* a combination of both 1D sliding and intersegmental transfers, along with the aid from other proteins such as helicase, polymerase etc. to a degree explains the apparent efficiency of the system.[177] In retrospect, synaptic joints formed with longer filaments with relatively greater number of nucleotide interactions must stay and eventually accumulate on the scaffold to circumvent energy required in dissociation and re-association with DNA, especially when ATP-hydrolysis is not needed during the homology search.[201] But surprisingly in the study by Forget et al synapses lasted only few seconds and only one such association was observed with λ DNA even for 168 nt long presynaptic filament.[11]

Although elegant, none of these methods offered direct observation of the intermediates formed during the homology search, giving insight as to how presynaptic filaments in vivo might sample genomic DNA confined in a crowded cellular environment. Despite being single molecule studies, such indirect measurements only observed a one-to-one 'yes or no' style interaction.[223,11,15]

A further advance was made by De Vlaminck et al, using dual magnetic and optical tweezers to examine the interaction forces occurring between a single nucleoprotein filament and DNA.[188] In this work, a variety of arrangements were explored, for example; the tethering of a DNA substrate between two optically trapped beads, whilst a RecA nucleoprotein filament tethered between the surface and a magnetically trapped bead is allowed to interact (figure 3.21).

This complicated setup enabled the authors to provide further evidence in support of the three dimensional freedom of the interacting DNA, suggesting that induced torsional stress, associated with the local under-winding and opening of the incoming dsDNA helix is critical to homology probing.[188,11] Furthermore, this also implies that multiple contacts can be made simultaneously, depending on the length of the nucleoprotein filament, as the incoming DNA does not require a free termini to probe from.

However, crucially this work provides evidence for two distinct binding events; a transient binding and a stable binding. Logically, this suggests the former relates to the initial residence of dsDNA in site II binding pocket identified by Chen et al.[159] Thus it supports the idea that short regions can be transiently probed, without the large energetic commitments.[160]

Interestingly, this work adds credence to recent work by Qi et al,[195] where transient binding is a necessary requirement for their proposed reductionist approach to a genomic wide homology search. Utilising a TIRF approach, they observed the residence times of dsDNA probes across a specially designed Rad51 nucleoprotein surface tethered curtains. Their results suggest, that homology searching occurs through the probing of 8 nt long homologous microdomains.[195]

This, the authors suggest, results in an efficiency increase as the entire genome does not require searching only a reduced set of sequences. Despite raising an interesting discussion, this concept is hard to reconcile in practice, as the nucleoprotein filament must still be required to sample a large sequence space in order to locate these microdomains of homology, around which further sequence homology is then probed for. However, that being said, there is plenty of evidence to suggest that initial recognition occurs in short sequences[192,193,194] and that this initial recognition exists in a transient state.[188]

Indeed, transient initial sequence probing is supported by molecular dynamics (MD) modelling.[224,221,161,160] This work has been made possible by the large resource of SAND, ET and XRD data produced for various proposed reaction intermediates.

One of the primary conclusions of such studies, is that the RecA provides little more than a passive supporting role in DNA strand exchange until the instigation of ATP hydrolysis at the end of

the mechanism. Thus acting to confine the initiating ssDNA and the invading dsDNA species proximally to allow base pair probing to occur through thermal fluctuations in a energetically favourable environment.[224,221,160]

Interestingly, this is in contrast to some of the earlier proposed mechanisms, which indicated RecA as an ATP dependant motor protein, performing a direct catalytic function.[7] Although this may not be a strictly redundant conclusion just yet, as ATP is shown to be crucial to full strand exchange, suggested to provide the final enthalpic "kick" that stabilises the new base pairing and dislodges the exiting DNA strand.[224]

3.4.3 High Speed AFM for Studying Biological Molecules

AFM offers many advantages for the observation of biological molecules which are observed without staining or fixation within native physiological environments. With respect to dynamics, time-resolved AFM investigations have been successfully carried out over the past two decades, offering critical insight into many areas of biological study. To date, many model systems have been observed in action, such as RNA polymerase,[225,226] bacteriophage Lambda Cro protein,[227] DNA photolyase,[228] nucleosome wrapping[229] and restriction enzymes[230] to name but a few.

Despite the success of these studies at revealing important information on the dynamics of biological molecules, the limited temporal resolutions of traditional AFM have impeded true real time observations. Biological reactions take place on the order of seconds and it is only recently that AFM technology has developed far enough to enable studies at multiple frames per second.

With the development of high-speed AFM over the last decade the ability to follow biological processes with sub-second resolution whilst maintaining single molecule spatial resolution is now becoming possible. One of the most astounding examples, work by the Ando group, demonstrates the observation of the hand-over-hand walking motion of the motor protein myosin V, confirming the long postulated mode of motion (figure 3.22).[231] Additional work by the same group has demonstrated the direct observation of the response of Bacteriorhodopsin to light[232] and the response of rotorless F1-ATPase to ATP.[233]

Furthermore, the Sugiyama group have demonstrated an array of dynamic studies of DNA structural transitions situated within DNA origami reference frames, including folding of G-quadruplex and identifying its intermediate states,[234,235,121] and the transition of B to Z form DNA.[126] Moreover, they have utilised the same system to observe directly the action of Holliday

junction resolving enzymes,[120] T7 RNA polymerase,[127] and the site specific recombination protein Cre.[236]

As such, AFM is reasoned to be the perfect tool for directly interrogating the interaction of NPF with dsDNA enabling direct insight into the homology searching mechanism.

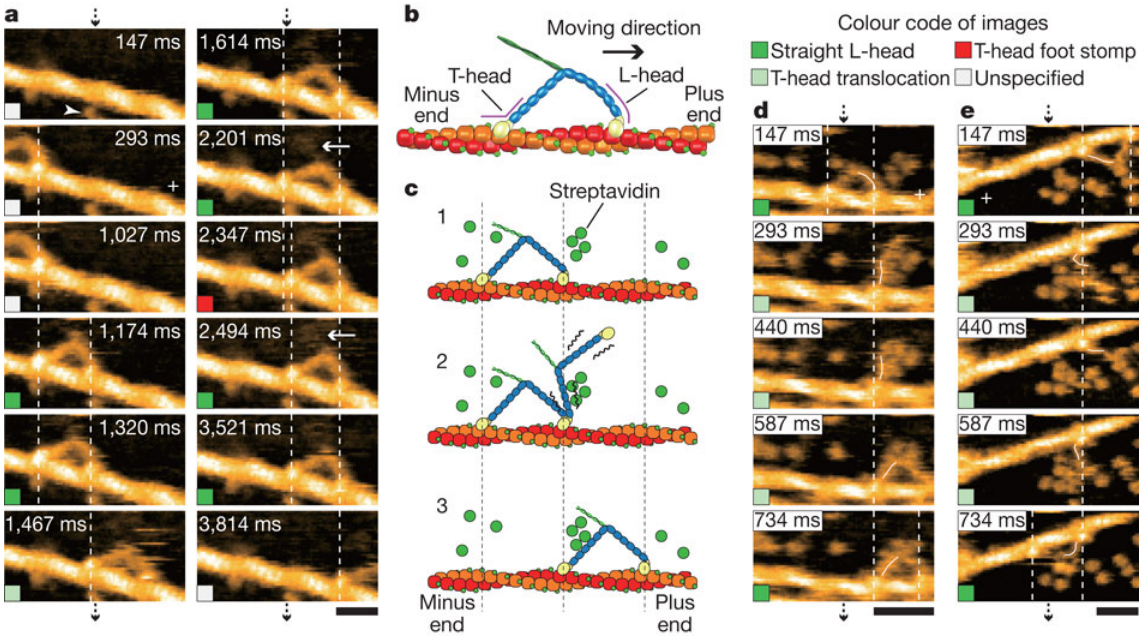

Figure 3.22: ***High speed AFM observation of Myosin V walking along actin.*** ***(a)*** *A sequence of HS-AFM images showing the progressive movement of Myosin V along an actin filament in the presence of 1 mM ATP. The arrowhead indicates a molecule of streptavidin which forms part of the sample preparation. The arrows indicate the coiled-coil tail of the Myosin V. The scale bar indicates 30 nm.* ***(b)*** *and* ***(c)*** *Schematic diagrams depicting the structure of Myosin V and walking mechanism, respectively.* ***(d)*** *and* ***(e)*** *A sequence of HS-AFM images depicting the hand-over-hand motion of Myosin V in 1 mM and 2 mM ATP, respectively. The scale bars in (d) and (e) indicate 50 nm and 30 nm, respectively. The centre of mass is indicated in (a), (d) and (e) with dashed lines and a plus denotes the positive end of the actin filament. The swinging lever of the motor protein is indicated by a white line. Image re-produced from Kodera et al.*[231]

3.5 Conclusions

This literature review has given a detailed examination of the extant developments in DNA nanotechnology and the current understanding of RecA, including examples of its use in bionanotechnology applications.

Since its inception,[1] the field of DNA nanotechnology has gone through some major revolutions: from the introduction of Motif structures;[43] the introduction of DNA origami;[2] to the creation

of wire-frame structures[80] and development of nano-canvases.[79] Despite their differences, these approaches all principally harness the high fidelity and predictable sequence structure of DNA to create novel nanoscale structures and arrange materials - both organic and inorganic - to develop functional devices. Furthermore, work to integrate these DNA-based devices together and with support surfaces has been highlighted.[3]

Interestingly, this section has only touched on a fraction of the rapidly expanding DNA nanotechnology field, discussing the structures and developments relevant to this thesis. It is important, therefore, to highlight that this text has not explored the associated developments in functional nanodevices and robots; including DNA tweezers, walkers and nanoscale assembly lines.[237] Neither has it delved into the associated field of DNA computation,[238] both of which are beyond the scope this thesis.

In light of this, it is becoming evident that as the field of DNA nanotechnology matures it is beginning to look for practical outlets in order to remain relevant. Regarding inorganic devices, the production and integration of nanoelectronic, nanophotonic devices or use as nanoscale masks for conventional photolithographic processing have all been demonstrated.[60] In contrast, DNA organisation of biomolecules is an incredibly promising route to novel biosensing applications, experimental platforms for the study of biomolecules and the targeted delivery of active agents in medicinal applications.[60,129]

With these applications in mind, this review further examined the recombination protein, RecA as a potential component for the hierarchical assembly of nano-electronics on these DNA templates. The current understanding of RecA structure and function was examined, with specific attention paid to the methods of study.

From the work discussed above, we can see that a true single molecule approach is required to confirm the proposed mechanisms adequately. Although many elegant studies have been carried out to date, providing information regarding the kinetics or DNA mechanical dependencies of these interactions, they fall short of providing clear experimental evidence in support of the insights provided by molecular dynamics simulations.

With respect to this, the approach taken in this project, using high speed AFM to observe the interaction in real time, has clear advantages over previous methods, not least in its potential ability to provide direct and real time observations of the interacting species. In the next section, the governing principles of the AFM and its applicability to the desired task are examined.

Chapter 4

A Technical Grounding to the Atomic Force Microscope

In this chapter, the underpinning theory of the AFM is explored in justification of its use for the observation of biological interactions and bionanotechnology devices. It is important to note that the text in this chapter is derived from a book chapter on the topic of the application of AFM to the study of bionanomaterials published by the author. For the original text please see Lee & Wälti.[239]

This chapter begins with a brief history of the microscope, including the development of the modes of operation. This is followed by a detailed description of the operational theory; including an examination of the tip-sample interaction forces, beam mechanics and the inherent resolution limits of the system, both spatial and temporal.

The chapter goes on to explore how these properties of the microscope enable the AFM to sense and control the application of forces with nano-scale spatial precision. This is discussed in relation to the extraction of nano-mechanical properties of soft-matter systems and how this exquisite force control enables access to higher spatial resolutions than were previously attainable.

Finally, recent developments which have taken place to increase the temporal resolutions, sufficient to follow the specific interactions of biological molecules are examined providing the theoretical grounding for the observations carried out in this thesis.

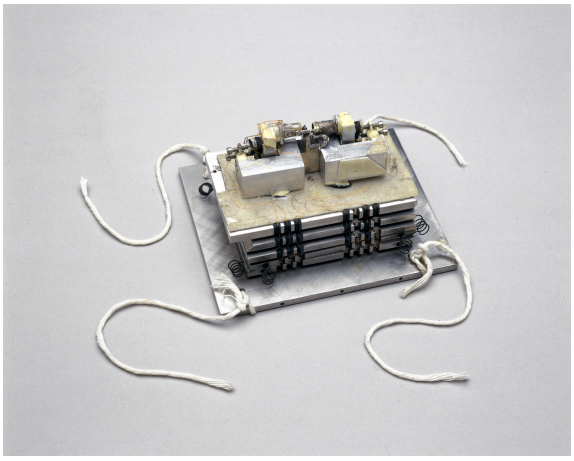

Figure 4.1: ***The first Atomic force microscope designed by C. F. Quate, G. Binning, and C. Gerber.*** *The core components of the microscope can be seen perched on top. All other the components, including the stacked metal plates, are providing the crucial vibration isolation of the instrument. The whole instrument was suspended from each corner using the attached cords for further isolation from vibration. Image reprinted with permission from Science Museum, Science & Society Picture Library, London.*

4.1 Brief Historical Perspective

The original AFM[210] as seen in figure 4.1, was based on the simple concept of moving a sharp tip at the end of a cantilever across a sample surface while recording the deflection of the cantilever. The cantilever was maintained in constant contact with the sample surface during imaging and this mode of imaging is nowadays referred to as contact mode (figure 4.2B).[210,240]

The AFM can be broken down into a few basic constituents, a cantilever and probe, a system for detecting motion of the cantilever and several piezoelectric elements arranged to coordinate movement of either the tip or sample in X, Y and Z. These elements can occur in various arrangements such as a sample scanning setup seen in figure 4.2 A.

In the original work, the deflections of the cantilever as a result of tracking the topology of the surface were measured by the tunnelling current occurring between the cantilever and an scanning tunnelling microscope (STM) tip positioned at a fixed location above.[210] The conceptual simplicity and the stunning imaging resolution that was achieved by this concept – lateral and vertical resolutions of 30 Å and 1 Å, respectively – led to a large variety of investigation being published in rapid succession, imaging a multitude of objects in air, liquid and under vacuum. The latter environments often resulted in significant enhancements to the spatial resolution due to removal of the capillary forces present in ambient environments owing to the accumulated water layers between the AFM tip and the surface.[241] This will be examined in more detail below. At this stage, AFM imaging was restricted mostly to hard surfaces while the examination of biological samples was limited because of the high lateral forces imposed by the tip during imaging, causing deformation and disruption of the substrate.

A first solution to this problem was introduced by Martin *et al* in 1987 in the form of a non-contact imaging mode.[242] Here, the cantilever was maintained within 1–10 nm from the surface and oscillated at its resonance frequency. Perturbations in the oscillation amplitude were used to form an image. In this imaging mode, the vertical and lateral forces imposed on the surface are negligible, making the imaging of delicate biological samples a possibility. However, this came at the price of significantly reduced spatial resolution (figure 4.2 D).

Modern AFM can be operated in a variety of different imaging modes, utilising different ways of how the probe interacts with the sample, taking into account the specific requirements of the surface to be imaged. In particular the introduction of dynamic imaging modes led to significant advances in the imaging of soft and in particular biological samples. Here, instead of maintaining a constant height as in non-contact mode and measuring the changes in amplitude as a result of the topography of the sample, the vertical position of the cantilever is adjusted dynamically such that a particular tip–sample interaction property, for example the amplitude of the resonant oscillations of the cantilever, is maintained. This is achieved using a feedback loop, and the output of the feedback loop is then used to form a topographical representation of the surface.

The most commonly used dynamic mode is an intermittent contact mode, commonly referred to as "tapping mode" (figure 4.2 C).[243] Again, the cantilever is excited at its resonance frequency and is subsequently brought into proximity with the surface. In contrast to non-contact mode, the cantilever tip lightly taps the surface at the extent of each oscillation, which leads to small distortion in the oscillation. Instead of maintaining a constant height and measuring the changes in amplitude as a result of topography, the amplitude is maintained and a feedback loop is used to adjust the vertical position of the cantilever relative to the sample surface via the actuation of a z piezo, to keep the distortion in oscillation constant. This method imposes higher vertical forces to the sample surface compared to non-contact mode imaging, but lateral forces remain low. Importantly, it provides greater spatial resolution over its true non-contact counterpart.[241] Shortly after the first demonstration in ambient environments, Hansma *et al* demonstrated tapping mode AFM in fluids in 1994,[244,245] which is arguably one of the major breakthroughs that opened up AFM techniques to the world of biological molecules. Additional enhancements to the spatial and especially the temporal resolutions of dynamic AFM modes have occurred over the last decade, increasing the applicability of the atomic force microscope in the study of biological molecules. As this thesis is concerned with the AFM imaging of biologically templated materials, we will focus mainly on dynamic imaging modes, although some of the theoretical aspects will also be applicable for general AFM operation.

Irrespective of the particular mode of imaging used, the accurate monitoring of the deflection of the cantilever is of key importance. As mentioned above, in the original work by Binning *et al*, the deflection of the cantilever was monitored by measuring the tunnelling current occurring between the cantilever and an STM tip positioned above.[210] An alternative detection method utilised piezoresistive layers applied directly to the top surface of the cantilever.[246] However, the most widely adopted method of cantilever detection was established in 1988 by Meyer *et al*.[247] Here, a laser beam is focused on the back of the cantilever. The reflected beam is then directed onto the centre of a four quadrant photodiode. For the non-deflected cantilever, half the laser beam now hits the upper and the other half the lower photodiode (figure 4.2 A). Upon deflection of the cantilever, this distribution is shifted and hence can be used to quantify the deflection. This optical system enables convenient and high-precision measurements of deflection as the small changes in deflection of the cantilever are amplified into large linear movements across the surface of the optical beam detector (OBD), allowing for small height features on the surface to be registered, according to:

$$\Delta Z \approx 2\frac{D}{L}\Delta z \tag{4.1}$$

where ΔZ is the change in position of the deflected laser beam on the OBD, Δz the change in z-position of the free end of the cantilever, D the cantilever–OBD distance, and L the cantilever length.

4.2 Background Theory

Despite the wide adoption of AFM, a solid understanding of the tip–sample interaction and cantilever dynamics has taken a long time to be established. Moreover, no single description thus far accounts for all imaging conditions. It is this limitation in theoretical understanding that has hampered rapid advancement in AFM techniques, especially with respect to minimising interaction forces while maintaining spatial resolution. However, recent advances enabled the development of advanced imaging regimes such as peak force tapping (PFT) which address these challenges (see section 4.4.2). The following discussion examines the physical interactions occurring between the tip and sample that underpin the formation of AFM images.

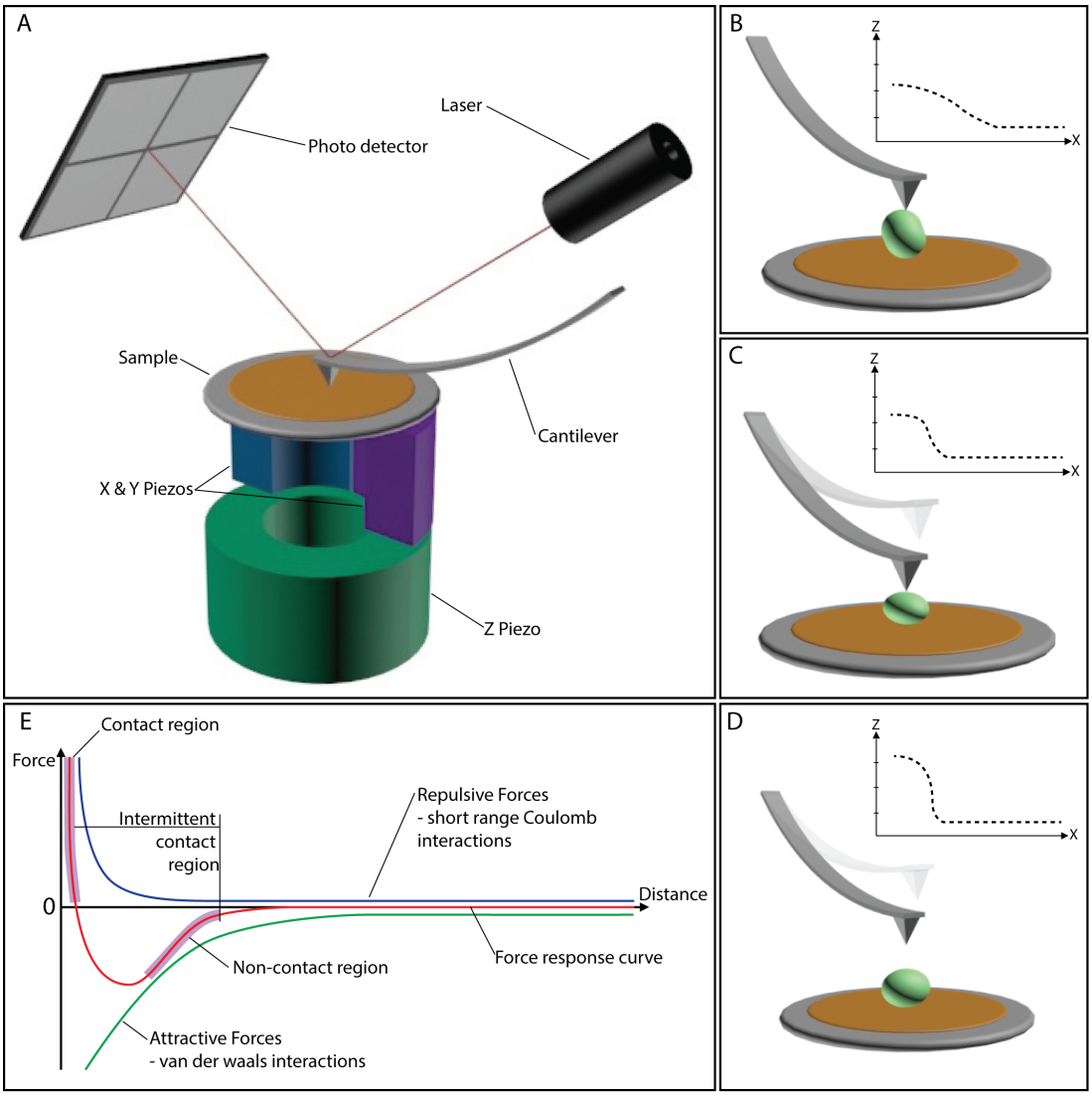

Figure 4.2: ***Schematic diagram depicting a simplified AFM setup.*** ***(A)*** *The main components of an AFM are depicted in a sample scanner arrangement, with the X,Y and Z piezoelectric elements actuating movement of the sample relative to the fixed position of the cantilever. Alternatively, a tip scanning setup can be used with the piezoelectric elements actuating the movement of the cantilever relative to the sample. A laser beam is reflected from the back of the cantilever to an optical beam detector. Minute cantilever deflections are amplified as linear movements of the laser across the optical beam detector. The conventional modes of cantilever actuation are depicted with associated image traces,* ***(B)*** *contact mode,* ***(C)*** *tapping mode and* ***(D)*** *non contact mode.* ***(E)*** *A force–distance curve depicting the force experienced by the cantilever upon approach to the surface. The imaging modes are indicated in their relevant region of the interaction curve.*

4.2.1 Tip–Sample Interactions

When considering the formation of an AFM image, one must consider the combination of attractive and repulsive forces that cumulatively act upon the tip to form the observable quantity (figure 4.3). These interaction forces are, however, difficult to disentangle and understanding each contribution in isolation is not trivial. The interaction between the tip of an AFM cantilever and a substrate is mediated by an array of electromagnetic forces, examined in brief below. Other more comprehensive reviews exist in the literature, for example by Garcia.[241]

First and foremost Van der Waals (VdW) forces must be considered, these are long-range attractive interactions between atoms and/or molecules arising from electric dipole interactions. In the case of tip–sample interaction, the sum of all the dipoles within both the sample and tip (figure 4.3 A) is of critical importance. Surfaces are rarely flat on the nanometre scale and hence the surface is approximated as an array of half-spheres. When considering the tip as a half-sphere a distance d away from the surface, the VdW force is given by:[248]

$$F_{VdW} = -\frac{H}{6d^2}\frac{RR_s}{R+R_s}\,, \tag{4.2}$$

where R_s is the effective radius of the surface roughness, R the radius of the tip, and H the Hamaker constant. As will be reviewed later, this view of a nanoscale surface defined by interconnected half-spheres, in part, defines the limit in spatial resolution achievable when imaging with another half-spherical object, i.e., the probe.

In addition, a range of repulsive forces, including those arising from Pauli exclusion and ionic repulsion, must be considered. However, these interactions can be described sufficiently well by phenomenological contact mechanics models, which take into account the load (applied force) and the elastic properties of the materials involved (both tip and surface). Hertz first described the contact and subsequent deformation of two spherical objects under an applied force in 1881.[249] Subsequent revised models by Derjaguin, Muller and Toporov (DMT)[250] and Johnson, Kendall and Roberts (JKR) elaborate on the relationships between deformation and load by additionally considering an adhesion force. Upon first encounter, the JKR and DMT models appear contradictory, however they describe opposing ends of the elastic modulus spectra.[251] Thus, the DMT model is relevant for low adhesion forces and stiff contact moduli when applied to small tip radii.[251] In contrast, the JKR model successfully describes large tip radii with a contact that is of low stiffness and has high adhesive forces. It is important to note that the former

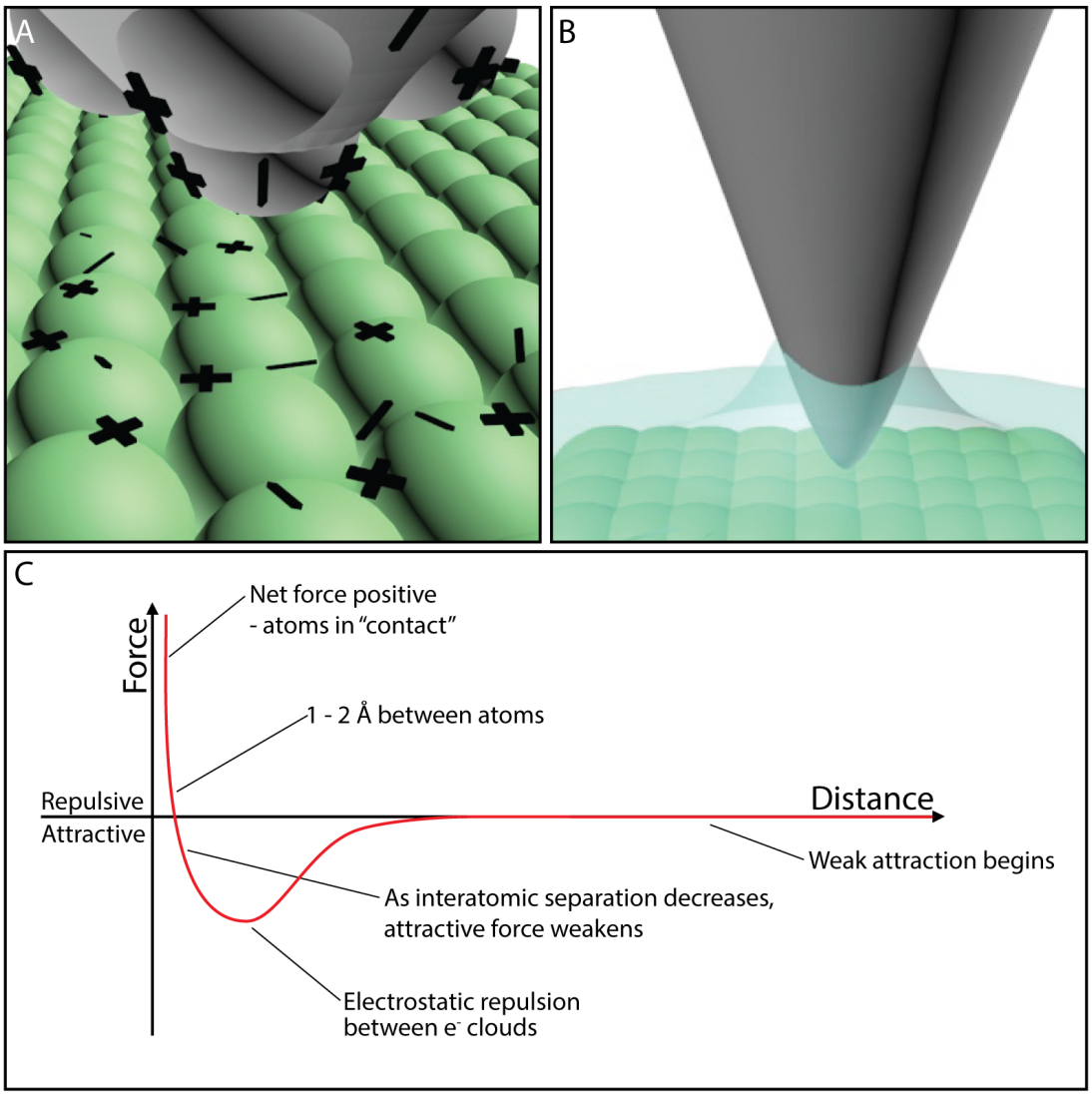

Figure 4.3: ***A montage of schematic diagrams depicting tip–sample interactions.*** ***(A)*** *A surface and tip are represented as ensembles of halfspheres with permanent and induced dipoles with the relative charges indicated. Van der Waals forces act between the dipoles of both the surface and the tip giving rise to attractive interactions.* ***(B)*** *A tip is depicted with an associated capillary neck as a consequence of an ambient water layer.* ***(C)*** *Interaction forces experienced between the tip and sample as a function of distance.*

considers adhesions forces acting outside of the contact area, whereas the latter does not.[241,251,252] In addition, these models also account for some non-conservative forces such as surface adhesion hysteresis and viscoelasticity.[241]

Operation Under Ambient Conditions

When operating the AFM under ambient conditions a water meniscus forms spontaneously between the tip and the sample surface, owing to capillary condensation (figure 4.3 B). The two principle radii of curvature of the meniscus, r_1 and r_2, are linked to the vapour pressure P of the liquid as follows:[253]

$$\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r_k} = \frac{R_g T}{\gamma_L V_m} \log \frac{P}{P_{sat}} \ , \tag{4.3}$$

where r_k is the Kelvin radius, R_g the gas constant, T the temperature, P_{sat} the vapour pressure at saturation, γ_L the surface tension of the liquid, and V_m the molar volume of the liquid. From this equation it can be seen that a meniscus will form when the tip reaches a distance from the surface equal to that of the Kelvin radius.

In general, analytical solutions of the tip–sample associated capillary force F_{cap} do not exist, but Israelachvili provides a solution in the approximation where the tip is modelled as a half-sphere and the surface is considered flat:[254]

$$F_{cap} = -\frac{4\pi R \gamma_L \cos\theta}{1 + \frac{d}{d_0}} \ . \tag{4.4}$$

Here, d is the distance between tip and surface, R the tip radius, d_0 is the height of the meniscus, and θ the contact angle at the surface. It is worth noting that the force reaches a maximum when the tip is in contact with the surface ($d = 0$). While the above equation is only valid for a specific approximation, the general cases are not dissimilar. It has been shown both numerically and experimentally that capillary force values can reach 100 nN, clearly becoming the dominating force in any ambient imaging regime when compared to other typical tip–sample interaction forces which range from high pN to low nN.

Under Liquid Medium

The prevalence of a dominating capillary force under ambient conditions can cause problems for stability and resolution, in particular when imaging soft specimens. This force can be eliminated altogether when the entire tip–sample interaction occurs within a liquid medium. However, other forces such as solvation and electrostatic double layer forces come into play owing to the presence of the solid–liquid interface at the tip surface.

When submerged in polar solvents, solid surfaces become charged through the absorption of ions or charged molecules, or by the ionisation of the surface atoms themselves. This excess charge at the surface is compensated by an accumulation of an equal number of opposite charges in the solution close to the surface. The thin layer of liquid containing the extra charges is generally referred to as the electrostatic double layer (EDL), which can extend into the solution from 1 nm to hundreds of nm, depending on the ionic strength of the solution. The EDL is highly dynamic, in constant battle between the entropy of ions becoming diffuse throughout the solution and the electrostatic attraction of the surface.

An EDL force is encountered when a second surface is introduced into the liquid and brought into close enough proximity to the first surface such that their respective EDLs overlap and hence interact. The EDL force is generally repulsive if both surfaces are either positively or negatively charged upon immersion in the polar solvent.

When imaging under liquid, the EDL force and the VdW force are usually dominating. These forces form the basis of the Derjaguin-Landau-Verwey-Overbeek (DLVO) theory, and an approximation of the resulting force in the case where the tip is modelled as a half-sphere and the surface is considered flat, was given by Butt *et al* for the case of $d \gg \lambda_D$, the Debye-length, i.e. the thickness of the EDL:[255]

$$F_{DLVO} = -\frac{4\pi R}{\epsilon\epsilon_0}\sigma_t\sigma_s\lambda_D \exp(\frac{-d}{\lambda_D}) - \frac{HR}{6d^2} \ , \tag{4.5}$$

where ϵ and ϵ_0 are the dielectric constants of the medium and the vacuum, and σ_t and σ_s are the surface charge of the tip and the sample, respectively. It is important to note that the thickness of the EDL varies strongly with the ionic strength of the medium, and hence the strength of the EDL force can be tuned accordingly. In doing so, access to high spatial resolutions is achievable.[256,257] Furthermore, an appropriate balance between a stable imaging environment, a suitable sample binding environment, and biological reaction conditions is required.[258]

In addition one must account for solvation and hydration forces that occur due to the confined nature of liquid molecules between the tip and sample surface.[259,260] It must also be considered that there is a dependency upon the mechanical and chemical properties of the substrate itself and its imaging environment.

4.2.2 Cantilever Dynamics

So far this text has only considered the complex interactions between the surface and the tip. However, one must be mindful of the role of the cantilever onto which the tip is mounted, in particular for dynamic mode imaging which is the main methodology for imaging soft matter. The cantilever can generally be modelled as a perturbed harmonic oscillator.[241]

The way in which a cantilever behaves when excited is defined by three characteristics: its resonant frequencies f_n, its quality factor Q, and its spring constant k. When considering a standard beam cantilever with the dimensions w, h and l, and only considering the most commonly used mode of oscillation, i.e. the transverse mode where the cantilever oscillates vertically, then the above characteristics are given by:[241]

$$f_n = \frac{1}{4\pi}\lambda_n^2 \frac{h}{l^2}\sqrt{\frac{E}{3\rho}} \,, \tag{4.6}$$

where E is the Young's modulus and ρ the mass density of the cantilever material. The coefficients λ_n are defined by $\cos\lambda_n \cosh\lambda_n = -1$, which leads to $\lambda_0 = 1.875$ and $\lambda_1 = 4.694$ for the lowest two resonances of the cantilever. The spring constant k is given by:

$$k = \frac{Ewh^3}{4l^3} \,. \tag{4.7}$$

The Q-factor describes any damping that affects a resonating cantilever and therefore is largely defined by the surrounding medium (air, liquid or vacuum). This dimensionless parameter can range from a few thousand in vacuum to a few hundred in air, and even lower in liquid, and effectively describes the "sharpness" of the principle resonance peak:

$$Q = \frac{f_0}{\Delta f} \,, \tag{4.8}$$

where Δf is the full width half maximum of the resonance peak. It is important to note that these equations are valid in vacuum, and that both, resonance frequencies and Q-factor can change considerably when the cantilever is oscillated in a medium.

4.2.3 Considerations in Fluid Environments

The desire to visualise biologically-templated materials, as well as eventually the actual assembly processes at the nano-scale, requires imaging in aqueous buffers which mimic their native

environments.[261,262] This further complicates cantilever dynamics owing to the changes in resonance frequencies as a result of the fluid environment (rather than vacuum) and viscous damping forces. Furthermore, when a cantilever is oscillated in a fluid environment there are two additional factors to consider. First, the interaction of the liquid with the cantilever, and second, the interaction at the sample–liquid–tip interface, where long-range forces such as EDL forces can become significant. However, it is important to remember that by imaging under liquid the capillary force, which can be dominating in ambient environments, is removed.[241,245]

For simplicity, this text only considers the case of imaging in a water environment, as EDLs formed by ions in aqueous buffers further complicate tip–sample interactions, and thus cantilever dynamics, and hence must be tuned appropriately depending on the system that is investigated. As mentioned above, in water there is a reduction in Q and f_n of the cantilever. For example when considering a beam cantilever, the primary resonance frequency is typically reduced by a factor of $3-5$ when operated in water compared with air. This decrease is caused by fluid boundary layers forming around the cantilever, which must be displaced during each oscillation cycle. Moreover, the viscous properties of the medium imparts hydrodynamic damping to the cantilever, which is responsible for the reduction of Q.[263,264]

As the oscillating cantilever is brought towards the surface, the fluid boundary layers form a greater proportion of the cantilever–sample separation distance, thus increasing the fluid shear, which in turn causes further damping. In most cases it is acceptable to neglect such shifts as these occur between the tip and the sample only, whilst the prevailing hydrodynamic effects act upon the entire cantilever body.[265] Additionally, higher harmonics can complicate the cantilever oscillations in liquid owing to the significantly reduced Q, while these were negligible in air.[266]

4.3 Spatial Resolution

Defining the spatial resolution of an AFM is different to that of conventional radiation-based microscopes. The AFM forms images of surfaces in three dimensions and hence requires the lateral (x and y) and vertical (z) resolution to be distinguishing separately from one and other. The main contributors that determine vertical spatial resolution are the mechanical and electrical noise in the feedback system, and thermal fluctuations of the cantilever. In contrast, factors such as the tip radius and aspect ratio, distance over which the surface forces decay and the compliance of the sample govern the lateral spatial resolution. Note that unlike with most radiation-based microscopy techniques, the mechanical properties of the sample play a defining role in the spatial

resolution obtainable, this is a factor which becomes important later (section 4.4).

4.3.1 Vertical Resolution Limits

The ability to distinguish small step heights is directly limited by the inherent noise of the imaging signal, either amplitude or frequency.[267] In the case of amplitude modulation, the most commonly used method, the ratio of the noise in the amplitude signal, δA, to the gradient of the amplitude versus tip–surface-distance defines the vertical noise inherent to the image:

$$\delta h_n = \frac{\delta A}{|dA/dd|} \ . \tag{4.9}$$

For soft samples such as biological molecules, the gradient is typically $0.2 - 0.5$, and hence δh_n is typically between $2\,\delta A$ and $5\,\delta A$.[267]

There are two major factors which contribute to noise in the cantilever signal, and these can be considered independent of each other. First, the cantilever experiences thermal instabilities and second, there is an inherent noise in the detector/feedback loop, dominated primarily by noise in the optical detection of the deflected laser beam. The latter has been reduced significantly over recent years through enhancements to the optical beam detector (OBD), improved laser sources, and small cantilever systems. As a result, the thermal instabilities of the cantilever are now the dominating noise-source of the cantilever in modern AFM systems.[268,259,269] The thermal vertical noise of the deflection signal can be approximated by:[270]

$$\delta h_{th} = \sqrt{\frac{4k_BTQB}{\pi f_0 k}} \ , \tag{4.10}$$

where k_B is the Boltzmann constant, B the noise bandwidth, and all other quantities as defined above. Provided that this is the dominating contribution and generally small in modern cantilevers, systems can be built which are sensitive even to the short-range forces proximal to the surface and therefore offer increased spatial resolution and force sensitivity, as the two are highly intertwined.[259]

4.3.2 Lateral Resolution Limits

The lateral resolution is governed by the typical feature sizes of both the tip and the sample, and the resulting AFM image is in fact a convolution of both shapes (figure 4.4). As a result, unless

the tip radius R is significantly smaller than the typical feature sizes of the sample, the lateral sizes of the imaged structures are generally overestimated by AFM. Furthermore, where typical surface features are smaller than the apex of the tip we can argue that the resulting AFM image is dominated by the tip apex rather than the sample surface. The lateral resolution, δl_r, can now be defined as the smallest distance at which adjacent features with a height difference of Δh on the sample can be resolved, and is given by:[271]

$$\delta l_r = \sqrt{2R}(\sqrt{\delta h_r} + \sqrt{\delta h_r + \Delta h}) \ , \tag{4.11}$$

where δh_r is the vertical resolution. In the context of this text, it is justified to set $\delta h_r \approx \delta h_n$. It is evident from the above equation that the lateral resolution does not only depend on the tip but also on the vertical resolution of the system, as well as the height of the features. For illustration, if we consider an AFM system with a typical vertical noise level of 0.1 nm and employing a tip with a radius of 5 nm, the lateral resolution δl_r when resolving peaks of identical height on the surface is 2 nm, while $\delta l_r \approx 4.3$ nm when imaging features with a height difference of 1 nm. Any features occurring closer than δl_r will be subjected to tip convolution artefacts (or dilation), this effect is represented clearly in figure 4.4.

In the above approximation, the tip apex is the only volume which is considered to interact with the surface. However, this is not always true, in particular when imaging surfaces where the height of the features is larger than the tip radius. In such cases, the side walls of the tip will also interact with the surface. To address this problem, the use of single walled carbon nanotubes (SWCNT) as AFM tips has been investigated owing to their excellent aspect ratio and tip radius.[272] However difficulties in preparing such tips impedes their wide adoption.

The above discussion of lateral resolution assumes non-compliance of both the tip and sample, i.e. neither sample nor tip deformation is taken into account. Although this is reasonable for solid samples, the elastic deformation of soft samples under an applied tip load will significantly alter both the vertical and lateral resolutions achievable.

When considered in relation to the Hertz model,[249] where applied force, F, tip radius, R, and the effective Young's Modulus of the sample, E_{eff}, are taken into account, the lateral resolution is found to be:

$$\delta l_r = 2\left(\frac{3RF}{4E_{\mathit{eff}}}\right)^{\frac{1}{3}} \ , \tag{4.12}$$

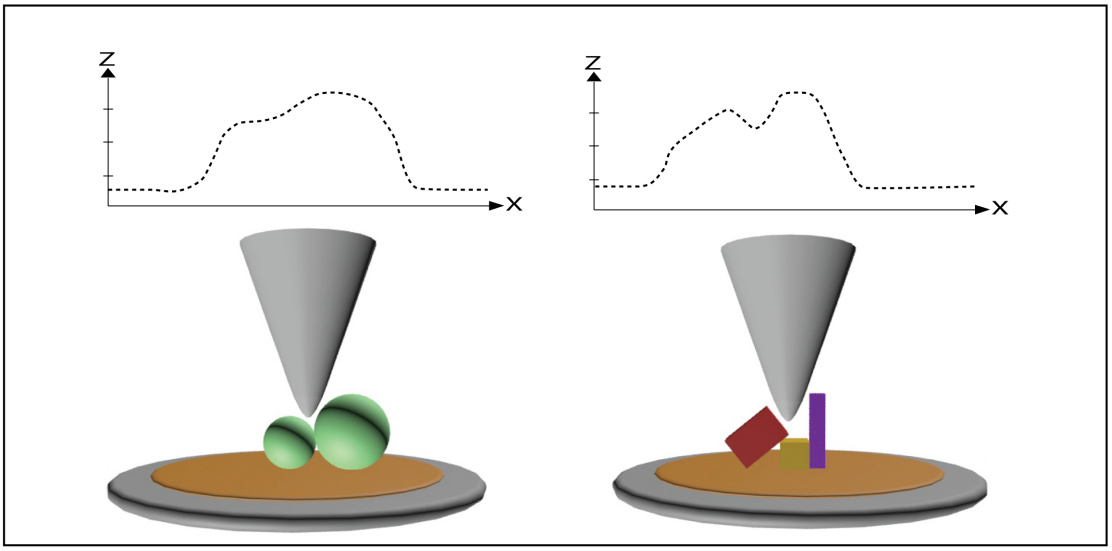

Figure 4.4: ***Schematic diagram illustrating the resolution limits when imaging objects with a probe of finite size.*** *Sample features smaller than the diameter of the probe cannot be resolved resulting in convolution artefacts, which represents a superimposition of the shape of the tip upon the sample.*

where the effective Young's modulus E_{eff} of the surface–tip interface is given by:

$$E_{eff} = \frac{1-\nu_s^2}{E_s} + \frac{1-\nu_t^2}{E_t} \; . \tag{4.13}$$

$E_{t,s}$ and $\nu_{t,s}$ are the Young's modulus and the Poisson ratio of the tip and the sample, respectively. This implies that forces below 1 nN and tip radii of around 1 nm are required to obtain sub-nanometre resolutions. More importantly, it indicates that such high resolutions are substantially easier to achieve on materials with high Young's moduli.[241,271] Interestingly, several groups have reported AFM images with molecular resolution of biomolecules by increasing E_{eff} by forming rigid structures (e.g. protein crystals).[256,273] In contrast alternative instrumental developments to circumnavigate the above challenges have been developed to control and reduce actively (peak force tapping, see section 4.4) and passively (torsional tapping[274,275,276]) interaction forces giving rise to higher resolutions.

4.4 Sensing and Controlling Forces

AFM employs a physical probe which interacts directly with the sample to gain the required information to form an image. This characteristic feature of AFM is, however, not limited to image formation. The typical forces experienced during imaging are of the order of pN to nN, and

therefore AFM lends itself to direct sensing of local interaction forces with high spatial and force resolutions. AFM probes can be utilised as nano-scale force sensors to investigate an external force acting upon it, and mechanical characteristics of a sample can be derived through the application of force. Properties such as deformation, energy dissipation, adhesion and moduli can be mapped with the same spatial resolution as available to topographical imaging.

Furthermore, by functionalising the AFM tip to introduce specific enhanced affinity for certain parts of the sample, AFM can be applied to the study of inter- and intra-molecular forces, including affinity mapping,[277] calculation of binding energies[278] and examination of the internal mechanics of proteins.[279]

4.4.1 Force Spectroscopy

The mechanical stability of biological systems is often critically linked with their biological activity, and thus being able to map nano-scale biological mechanics is of great importance in understanding how form and function are linked, and in this endeavour AFM is undoubtedly unsurpassed.

Considering the AFM cantilever as a Hookian spring, the interaction force between the AFM tip and a sample can be given as:

$$F = -k\Delta z \ , \tag{4.14}$$

where k and Δz are the cantilever spring constant and deflection, respectively.

Mechanical properties can be extracted by driving the cantilever along the z-axis from a position far away into the sample surface (and back) while monitoring the cantilever deflection signal. This is typically referred to as a force-distance curve, however, not the tip–sample distance but only the position of the z-piezo (Z) is known, which represents the rest-position of the cantilever. The tip–sample separation distance d is affected by the deflection of the cantilever and the elastic properties of the sample. Therefore, d is given by:

$$d = Z - (\Delta z + \delta_s) \ , \tag{4.15}$$

where δ_s is the sample deformation. Hence, the recorded curve contains elastic contributions from the cantilever and the sample, as well as contributions from the tip–sample interactions,[280]

including both long-range attractive and short-range repulsive (mechanical contact) forces.

Because of the elastic nature of the cantilever – and in many cases, the sample itself – discontinuities between the approach and the retract parts of the curve occur when considering the attractive interactions. The hysteresis mainly results from the snap-to-contact dominated by the formation of a capillary neck in ambient conditions, but also includes VdW and hydrophobic contributions. In general, more detailed information on the attractive part of the tip–sample interaction can be obtained upon retraction.

When considering the repulsive contributions, the spring constant of the cantilever plays a major part. Mechanical deformation of the sample occurs as the cantilever is lowered and the tip starts to interact with the sample surface until the cantilever begins to deflect due to the resistive force of the surface. Therefore, it is important to match the compliance of the cantilever and the sample to avoid damage to the sample and optimise force sensitivity, and this is of particular concern where biological samples are investigated. In addition, the tip geometry must also be taken into account, for example when investigating biological membranes to image living cells, a sharp, high aspect ratio tip will lead to high loading forces across a small surface area, and hence can lead to the rupturing of the membranes.

4.4.2 Mapping Mechanical Properties

Force curves are not limited to single points, and can in fact be collected across an entire surface analogous to, for example, imaging in tapping mode where height information is collected for each point across a surface. This process of mapping the mechanical properties across an entire surface is known as force-volume mapping. Traditional AFM instruments can collect force curves at a rate of up to 10 Hz, and typically each curve is treated independent of the previous one, which can lead to overshooting and increased sample load.

This approach of measuring force at each point across the sample-surface has recently been adapted and developed further to enable ultra-high resolution imaging. In this approach, generally referred rapid force-curve-based imaging (figure 4.5), the applied force is continuously analysed and the z-position of the cantilever is adjusted via a fast feedback loop such that the maximum force applied to the sample remains constant. Note that this effectively decouples the resonance of the cantilever from its response by oscillating at a set low frequency, typically fixed at 2 kHz.[281,282] Furthermore, the z piezo is actuated with a sinusoidal waveform, contrary to traditional force-distance curves which are derived using linear ramps. This results in the tip velocity approaching

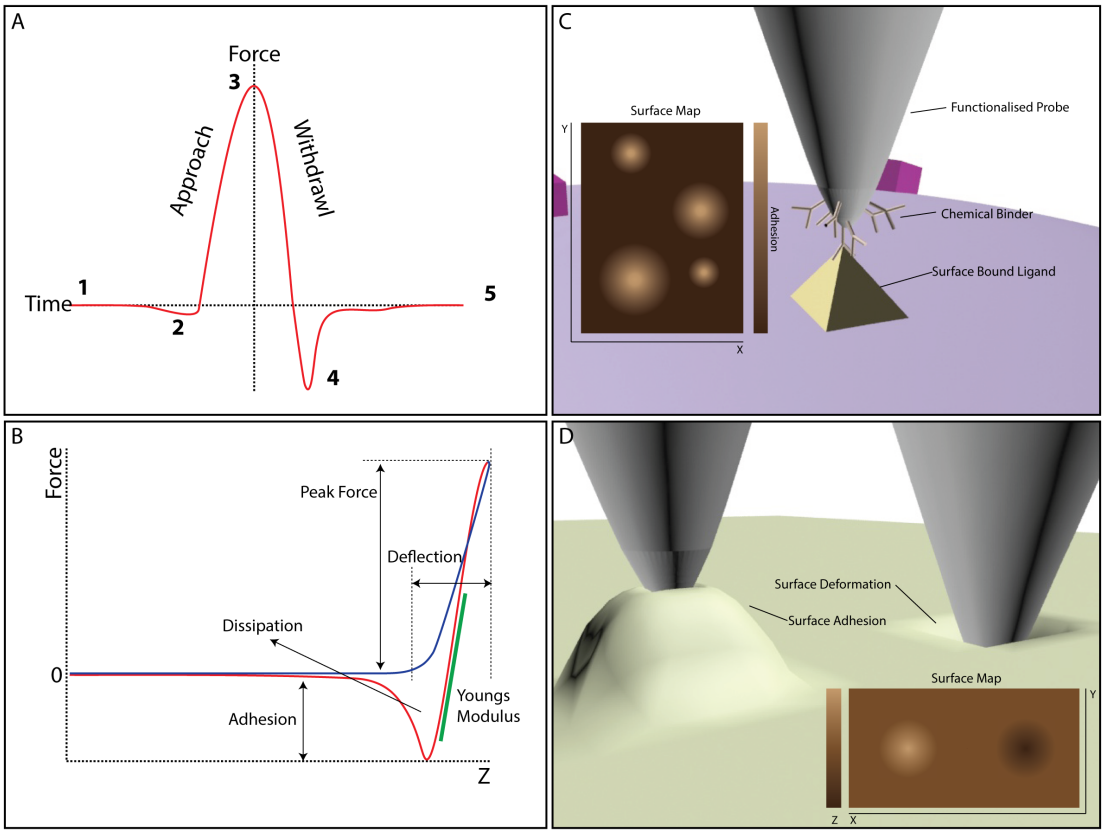

Figure 4.5: ***The Rapid force-curve-based imaging regime and applications*** ***(A)*** *The typical force–time profile of a Rapid force-curve-based imaging method. As the tip–sample distance is decreased, the cantilever goes from being at rest (point 1) to sensing long-range attractive forces (point 2). Upon further reduction of the separation, the cantilever experience mechanical loading until a set threshold is reached (point 3). An adhesion force is experienced by the cantilever upon retraction (point 4), and when overcome by the restoring force of the cantilever, the system returns to its rest (point 5). An associated force–distance profile is depicted* ***(B)*** *showing both the approach (blue line) and retract (red line) curves. Some characteristic quantities that can be derived from the interaction are indicated. The sensitivity of this method has applications in recognition mapping* ***(C)*** *and nano-mechanical mapping* ***(D)****.*

zero at the point of contact, enabling a more controlled interaction.[282,283]

Owing to the increased speed of modern controller electronics, the system is able to modulate the peak-force applied to the sample surface per oscillation cycle to a self-adjusting set-point at which sample deformation begins. In practice, the probe is able to interact with the sample with forces of tens of pN for merely 100 μs, with the system deriving multiple force-distance curves per pixel at a typical rate of 200 Hz. In addition, the signal-to-noise is drastically increased by algorithmically removing any parasitic components from the deflection signal at each surface contact.[282]

This approach of direct force control enables imaging at very low interaction forces and hence

leads to increased spatial resolution compared to traditional modes, in particular for soft polymer and biological samples. In addition, rapid force-curve-based imaging avoids excessive sample deformation,[284,285] minimises tip wear,[283] and maintains stable imaging as the set-points are not susceptible to drift, which is a particular benefit when imaging in fluid. A further important benefit of this mode is the fact that the mechanical properties are extracted at each pixel which enables qualitative analysis of the samples local characteristics (figure 4.5 C).[281,284]

Other variations of force mapping of biological samples includes the direct investigations of ligand-receptor interactions.[286] An early example of this method is discussed by Gaub *et al*, measuring the interaction forces of individual biotin–avidin complexes[287,288] and DNA hybridisation.[287] When coupled with force-volume or PFT, this method can be used to examine the spatial distribution of specific receptors, typically across a biologically active surface (figure 4.5 D).[277,278]

As discussed, the application of small forces gives access to superior spatial resolutions upon fragile samples, in particular those of biological nature. It is now possible to image bionanomaterials with sufficient resolution to probe their internal structures and identify perturbations as a consequence of hierarchical interactions such as the work of Pyne et al demonstrating variations of the helical pitch of DNA. This high spatial resolution offers insight into the exact role each biological component plays in a particular assembly, crucial information when attempting to manipulate their interaction to form bionanomaterials.

4.5 Temporal Resolution

4.5.1 Theoretical limits

It can be argued that one of the great limitations of AFM for the study of biological systems until very recently was its lack of ability to image at sufficient speed to monitor biological processes in real time. However, in the last few years, a number of instruments have become available which make this possible. Before these recent advances are reviewed, this section briefly considers the theoretical aspects governing temporal resolution of dynamic AFM under fluid.

The highest possible imaging rate is fundamentally defined by the feedback bandwidth of the control loop that maintains the tapping force during a scan, with further limitations implied by the sample itself.[262,289,290,291,292,293,294]

The image acquisition rate, r_{aq}, is given by:

$$r_{aq} = \frac{v_t}{2An} \ , \tag{4.16}$$

where v_t is the velocity of the tip (x axis), A the scan area, and n the density of scan lines (y axis). If we consider a sample whose typical features are distributed along the x axis with a periodicity η, the bandwidth ν_B of the z axis feedback loop must be sufficient to trace the sample surface, i.e.:

$$\nu_B \geq \frac{v_t}{\eta} = \frac{2r_{aq}An}{\eta} \ , \tag{4.17}$$

and hence

$$r_{aq} \leq \frac{\nu_B \eta}{2An} \ . \tag{4.18}$$

Provided that the feedback loop acquires a phase delay of $\frac{\pi}{4}$ at the bandwidth frequency, the time domain delay resulting from the feedback loop is at least $\Delta\tau = \frac{1}{8\nu_B}$. Most AFM systems operate in closed-loop feedback configuration, and it is reasonable to assume that the total time delay doubles as a result of the closed loop,[262] i.e. $\Delta\tau_{cl} \approx 2\Delta\tau$, and hence

$$\nu_B \leq \frac{1}{16\Delta\tau_{total}} \ , \tag{4.19}$$

where $\Delta\tau_{total}$ is the total time delay in the feedback which is equal or bigger than $\Delta\tau_{cl}$. Hence, the maximum acquisition rate is given by:

$$r_{aq}^{max} = \frac{\eta}{32\Delta\tau_{total}An} \ . \tag{4.20}$$

For illustration, in order to achieve an acquisition rate of 5 frames per second at a scan line density of 10 nm^{-1} for a sample of 50 nm x 50 nm with a typical feature periodicity of 1 nm, the total time delay in the feedback loop must not exceed 0.25 μs, and a minimum feedback bandwidth of 250 kHz must be achieved.

4.5.2 Instrumentation Perspective

Early efforts to increase the imaging speed of conventional AFM focused on contact mode techniques. It was identified that the mechanical response of the microscope is limited by the time delay in the piezo tube that actuates displacements in x, y and z. In order to increase the bandwidth limitations of these piezo stacks, feedback actuation was attempted by the deposition of piezoelectric films directly on the cantilevers.[295] Subsequent developments saw the attempted use of these integrated piezoelectric films in dynamic modes.[296,297]

Through these methods it was possible to increase the imaging speed by an order of magnitude, although it was limited by complicated signals that arose from integrating the piezoelectric films into the full length of the cantilever itself.[240] Furthermore, limitations in fabrication sizes of such cantilevers resulted in high spring constants and poor resonance frequencies,[298,299] which were further compounded in liquid.[300] Efforts to develop large arrays of individually addressable cantilevers were simultaneously attempted.

Further progress in fast-scan AFM was achieved by Miles *et al*[301,302] by developing novel approaches of constant height modes. In their method, a tuning fork, resonating at 100 kHz, was employed as the x scanner, enabling frame rates of up to 14 frames per second in ambient conditions when imaging semi-crystalline polymer samples.

Ando's and Hansma's groups simultaneously reported attempts to increase the feedback bandwidth of dynamic modes. A flurry of inventions followed, including small cantilevers with dimensions of less than 10 μm and corresponding optical detection systems to utilise these small cantilevers[245,303,304,305], active damping z scanners,[306,307] dynamic PID controllers,[308] and fast data acquisition systems.[309,310,311]

4.6 Conclusions

Since its invention, the AFM has proven to be an extremely powerful tool for the characterisation of materials at the nano-scale. The exceptional versatility of this form of microscopy has made it a key tool for the interrogation of biological molecules. Although the operating principles of most if not all aspects of AFM are conceptually straight-forward, the theoretical foundations for understanding the various different interactions involved have proven to be much more demanding to establish. In this chapter, the basic underlying theory of dynamic AFM modes has been examined.

The interrogation of a surface by a sharp tip in mechanical contact with that surface means that the spatial resolution is directly governed by the physical properties of both the tip and the sample, as well as being a function of the force applied. As a result achieving sub-molecular resolutions is challenging when imaging soft samples such as biological molecules and biologically templated materials, which have a tendency to deform under applied loads. Recent developments in rapid-force-curve have enabled exquisite control over the applied interaction forces enabling access to spatial resolutions previously unobtainable on biological molecules in isolation.

Furthermore, as AFM collect data in a serial fashion, the temporal resolution is governed principally by the response time of the mechanical and electrical components of the instrument's feedback loop. A number of key advancements to the temporal resolutions have been highlighted in this chapter, indicating substantial developments in AFM instrumentation. These recent development in HS-AFM is beginning to allow the interrogation of biological mechanisms over relevant time scales, which forms the basis of the work presented in this thesis.

Given theses advances in the principles resolutions of the the AFM it is justified as the preferential tool for interrogating the intricate structure and direct real time observations of the RecA mediated homology searching mechanism. The next chapter details the materials and methods used in this thesis in order to study the mechanism in question.

This page is intentionally left blank.

Chapter 5

Methods and Techniques

This chapter describes the core materials and techniques utilised throughout the thesis. This section aims to provide the base protocol, deviations from which are described where relevant in the thesis.

5.1 Materials

5.1.1 Protein and Chemicals

RecA protein (E. coli) was purchased from New England Biolabs (NEB), Inc. (Ipswich, MA, USA) at a concentration of 2 mg/ml in 20 mM Tris(hydroxymethyl)aminomethane - Hydrochloride (Tris-HCl) (pH 7.5), 1 mM Dithiothreitol (DTT), 0.1 mM Ethylenediaminetetraacetic acid (EDTA), and 50% glycerol and was utilised without further purification.

Adenosine 50-[-thio]triphosphate tetralithium (ATPγS) salt, which was made up to a final concentration of 5 mM in deionized water; Magnesium Acetate (MgAc), that was made up to 200 mM and Tris-Acetate (Tris-Ac), that was made up to 300 mM, pH 7.4 were all purchased from Sigma Aldrich (St. Louis, MO, USA).

5.1.2 DNA Fragments

All DNA oligonucleotides utilised in this thesis - including; PCR primers, patterning oligonucleotides and origami staples - were synthesised by Integrated DNA Technologies

(IDT) (Coralville, IA, USA) and purified by desalting. The nucleotide sequences for all the oligonucleotides used in this thesis are given in appendix B.

All the dsDNA templates were amplified using PCR from either a modified pGEM-T plasmid vector,[312] in the case of the 890 bp fragment or λ bacteriophage DNA, in the case of the 3.5 kbp fragment.

5.1.3 Buffers

Tris-Mg buffer

Concentration	Constituent	Description
10 mM	Tris-Acetate, pH 7.4	HS-AFM buffer containing minimum
10 mM	Mg-Acetate	Mg^{2+}concentration for RecA reactions.

5.2 Methods

5.2.1 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR) is a technique in molecular biology used to amplify the concentration of a particular DNA sequence over several orders of magnitude.

PCR utilises thermostable DNA polymerase enzymes, which catalyse the formation of phoso-di-ester bonds between adjacent nucleotides according to an ssDNA template. PCR proceeds as a one-pot reaction through a thermocycle to "melt" the dsDNA template, anneal short oligonucleotide primers flanking the target sequence and enzymatically – using the polymerase enzyme – extend the daughter sequences.

A typical PCR reaction consists of three-stages usually cycled up to 30 times. The product concentration increases in accordance with: $mx2^{n-1}$, where m is the original template concentration and n the number of cycles. The different stages of the PCR cycle are described below (figure 5.1):

Denaturation: The reaction is cycled to 95°C for 30 s to denature the dsDNA generating a ssDNA templates. It is important to note, a longer initial denaturation stage is typically carried to ensure all the DNA is in the form of ssDNA and is available for primer annealing.

Annealing: The reaction is cooled to 5°C below the T_m of the pair of primers, typically 50 - 65°C for 20 - 40 s. This stage allows the individual primers to hybridise at the complementary 5'

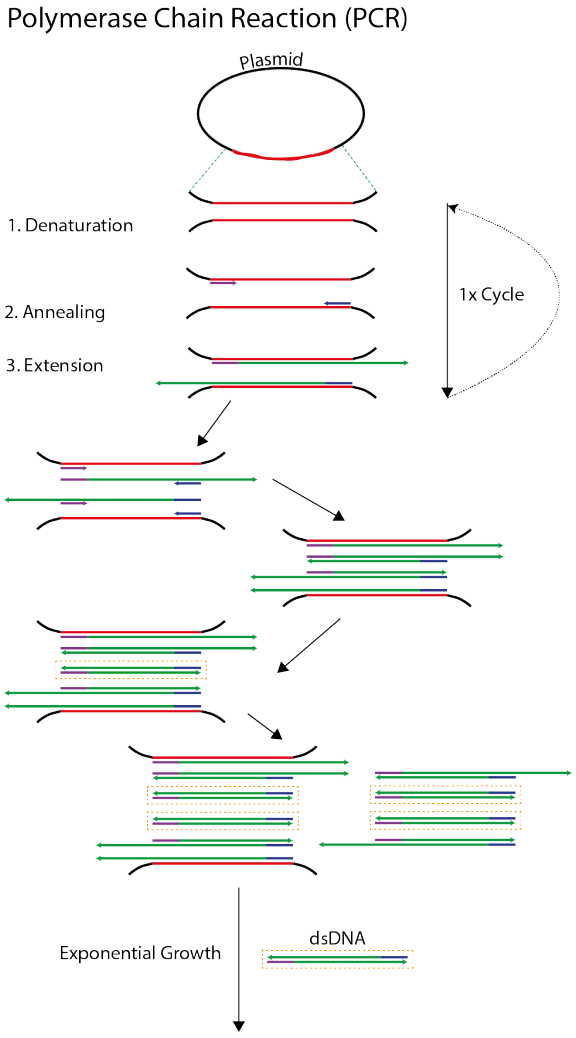

Figure 5.1: ***A schematic diagram depicting the Polymerase Chain Reaction (PCR).*** *A typical PCR reaction is cycled between; denaturation* ***(1.)****, primer annealing* ***(2.)*** *and polymerase extension* ***(3.)****. The target product dsDNA is amplified exponentially through successive cycles. The templates are indicated in red and the upper and lower primers in blue and purple, respectively. The predominant products are highlighted (encircled in an orange dotted box).*

termini of the target sequence on the ssDNA template.

Extension: The reaction is heated to 68 - 72$°C$, depending upon the working temperature of the polymerase used, allowing the extension of the daughter strand from the bound primer in the 5' - 3' direction, creating a new dsDNA molecule. The duration of this step depends upon the length of the target sequence being produced, with a typical rate of a 1 $kbp\ min^{-1}$. A final extension stage after the final cycle ensures all products are dsDNA.

In this thesis, PCR amplification was used to produce the 890 bp and 3.5 kbp dsDNA templates. A typical reaction consisted of; 1 μl of template DNA (10 ng/μl), 1μl each of the upper and lower primers (10 μM), 10μl of Long Amp Taq polymerase buffer (5X), 4 μl of mixed dNTPs (10 mM) and 1 μl of Long Amp Taq polymerase. The total reaction volume was adjusted to 25 μl with dH_2O.

5.2.2 Trailing End PCR

An adaptation on the standard PCR discussed above enables the inclusion of new terminal sequences into the product dsDNA compared to the original dsDNA template (figure 5.2).

This requires primers with distinct domains and an associated phasing of the PCR cycle to accommodate changes in annealing temperatures. The new sequence is included as a 5' flanking region (figure 5.2, green). The 3' portion of the primer contains sequence complementary to the template - as with standard PCR.

In the first phase the annealing temperature is optimised for the initial complementary region of the primer (figure 5.2, blue) - typically $10°C$ lower than the full primer. Extension of the primer results in the inclusion of the 5' flanking sequence, which is not annealed to the original ssDNA (figure 5.2 phase 1). This phase occurs over 5 cycles, resulting in a linear amplification of a new heterogeneous ssDNA sequence.

In the second phase, the annealing temperature is optimised for the full primer sequence including the newly incorporated 5' flanking sequence. This is conducted for 25 cycles and is equivalent to standard PCR, amplifying the intended product exponentially. Reactions were setup as described above.

5.2.3 DNA Purification and Quantification

All DNA products were purified to remove residual reaction components, including enzymes, Deoxynucleotide Triphosphates (dNTPs) and to exchange into a storage buffer. Purification was conducted using spin columns from either Roche (Roche Diagnostic Ltd, West Sussex, UK) or Qiagen (Manchester, UK); High Pure PCR Product Purification Kit or a QAIquick PCR purification kit, respectively. These systems make use of DNA binding to silica surfaces in the presence of chaotrophic agents and high ionic buffers. Despite differences in proprietary chemistries both provide comparable efficiencies and were thus utilised interchangeably.

Briefly, dsDNA samples were bound, washed and eluted from the columns utilising the provided buffers, as per the respective protocols. All spin steps were carried out in a tabletop microcentrifuge (Eppendorf) at 13,000 rpm at room temperature. All products were eluted in a typical volume of 30 μl of respective elution buffer, containing Tris-EDTA. Typically, for higher product yields, two PCR samples were concentrated in same column, processed and eluted together.

Subsequent species quantification was carried out using a Nanodrop 2000 UV spectrophotometer (Thermo scientific, Wilmington, DE, USA). The pedestal was rinsed with dH_2O prior to use and blanked against the respective elution buffer. 1 μl of sample was utilised to quantify the DNA yield. Finally the samples are pooled and the concentration averaged for the final stock.

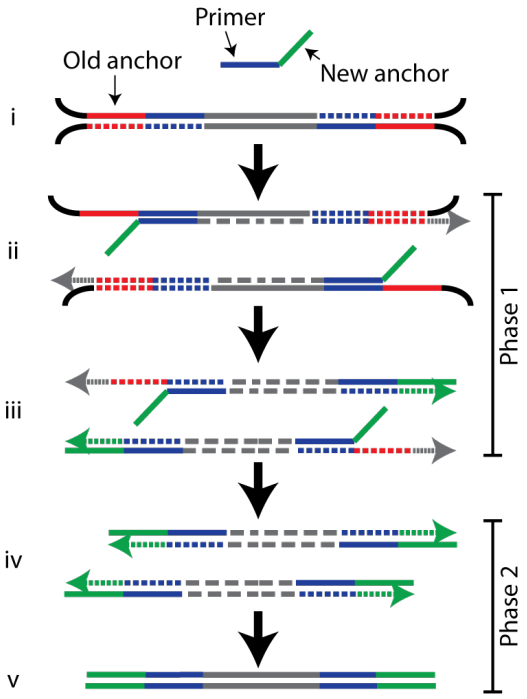

Figure 5.2: ***A schematic diagram depicting the trailing end PCR.*** *Primers containing separate sequence domains are utilised. The 3' region is complementary to the original template (blue). The 5' region contains a news sequence (green). The PCR is phased to optimise annealing temperatures to the binding of the 3' region only initially and the full primer later. The sequence being replaced is highlighted in red and the new sequence in green. Complementary regions are shown in the same colour.*

5.2.4 Analytical Gel Electrophoresis

Gel electrophoresis is a common method in biology for separation and isolation of populations of biomolecules based upon their molecular mass, charge and conformation. This method separates biomolecules such as DNA, RNA and proteins by sieving through a porous gel matrix under an applied electromotive force.

The gel matrix is submerged in a running buffer - either Tris-Acetate-EDTA (TAE) or Tris-Borate-EDTA (TBE) - between a cathode and an anode, such that electrical connection is made through the gel. An applied electric field, typically a constant voltage, imparts an electromotive force upon the charged biomolecules causing them to migrate through the sieving matrix towards the cathode or anode depending upon the net charge it holds.

A molecules electrophoretic mobility, μ, is the velocity per unit of field strength, as given by:

$$\mu = v/E = d/Et \tag{5.1}$$

where d is distance (cm), t is time (s), v is the migration velocity and E is the field strength. Therefore, electrophoretic mobility is expressed as $cm^2V^{-1}s^{-1}$.

There are two common types of gel matrices utilised in electropheretic systems: polyacrylamide and agarose. Typically, agarose gels are found to offer a greater range of separation due to its large pore size, able to separate DNA molecules from 50 Base pair (bp) to 20 Kilobase pair (Kbp) in size.[313] In contrast, polyacrylamide typically has a much greater resolving power due to its smaller pore sizes and as a consequence are utilised to separate DNA molecules that differ in several tens of bp or less.

A typical analytical gel electrophoresis process consists several distinct steps;

- Casting the gel matrix;
- Loading samples (in a glycerol loading dye);
- Separation at a constant Voltage;
- Visualisation of the migration pattern.

It is important to note that visualisation is conducted by staining the DNA with intercalating dyes - such as ethidium bromide or the SYBR family of dyes - which are excited at 302 Nanometre (nm) UV-light. These are sensitive down to picogram DNA concentrations. Note that agarose gels are typically impregnated with the dye while polyacrylamide gels are post-stained.[313]

In this thesis, all PCR products - i.e., 890 bp and 3.5 kbp - were assessed with agarose gels, typically 1%. Typically 100 ng of product was loaded and run along side a mixture of DNA fragments of defined lengths – commonly referred to as a ladder – for verification.

All other gels presented in this thesis were polyacrylamide, where the percentages are noted at the relevant place in the text or figure caption. Typically, 10 ng of sample was loaded alongside a 10 or 100 bp ladder, noted where relevant. Separation was typically conducted at 80 V for 90 - 240 minutes.

5.2.5 Generation of Single Stranded Overhangs by Nicking and Competing

This method was used to generate ssDNA overhangs on dsDNA templates, such as those required for inclusion into the DF2L origami frame. The generation of these ends was mediated in two steps; nicking of the dsDNA backbone and the removal of the unwanted ssDNA region

by denaturing the template DNA. The short ssDNA region was sequestered with an excess of complementary oligonucleotide during re-annealing of the DNA template.

Nicking

Recognition sites for the nicking enzyme nt.Bst.NBI were designed into the templates at the relevant places in the opposing termini. A typical reaction consisted of 10 μl of dsDNA template (@ 50 - 100 ng/μl), 5 μl nt.Bst.NBI (@ 10 Units/μl), 5 μl NEB Buffer 3.1 (@ 10 x) made up to a total volume of 50 μl with dH_2O. This was incubated for 5 hours at 55$^\circ C$.

Competing

Competing oligonucleotides were introduced to 5 μl of nicked templates at a 100 fold excess. The reaction mixture was made upto 50 μl including NEB buffer 3.1 (@ 10 x) and dH_2O. The mixtures were denatured at 80$^\circ C$ for 20 minutes and allowed to cool to room temperature in a controlled gradient of -1$^\circ C$ min^{-1}. The final products were purified as described in section 5.2.3.

5.2.6 RecA Patterning Reaction

RecA patterning complexes were formed as established in Sharma et al[5,6], with minor deviations described where relevant. The RecA mediated patterning reaction is orchestrated in two phases; the formation of a NPF on ssDNA (figure 5.3 I) and the interaction of an NPF to form a region of triple stranded DNA (figure 5.3 II).

RecA Nucleoprotein Filament Formation

The amount of RecA (R_V) required to polymerise on a given oligonucleotide is a function of its length (O_L), its concentration (O_C) and the fact that RecA binds 3 nt per monomer. Hence;

$$R_V = \frac{(O_L/3)O_C}{R_C} \tag{5.2}$$

where, R_C is the concentration of the RecA stock.

A typical reaction for a 30 nt oligo consisted of 1 μl ssDNA (@ 10 μl), 3.8 μl RecA (52.3 pmol/μl), 1 μl ATPγS (5 mM), 1 μl Tris-Acetate (200 mM, pH 7.4) and 1 μl Mg-Acetate (20 mM). The final reaction volume was made up to 10 μl with dH_20 and incubated at 37°C for 15 minutes.

Synaptic Joint Formation with dsDNA

Freshly formed NPFs were introduced to the relevant template dsDNA in a specific ratio and further incubated at 37°C for up to 60 minutes (figure 5.3 II). Both the ratios and incubation times vary depending on the experiment and are discussed where relevant in the text. Typically, where a 3:1 (NPF:dsDNA) is utilised; 1 μl NPF (@ 1 pmol/μl) was added to 1 μl dsDNA template (@ 0.346 pmol/μl), 1 μl Tris-Ac (200 mM, pH 7.4) and 1 μl Mg-Ac (@ 200 mM). The final reaction volume was made up to 10 μl with dH_20.

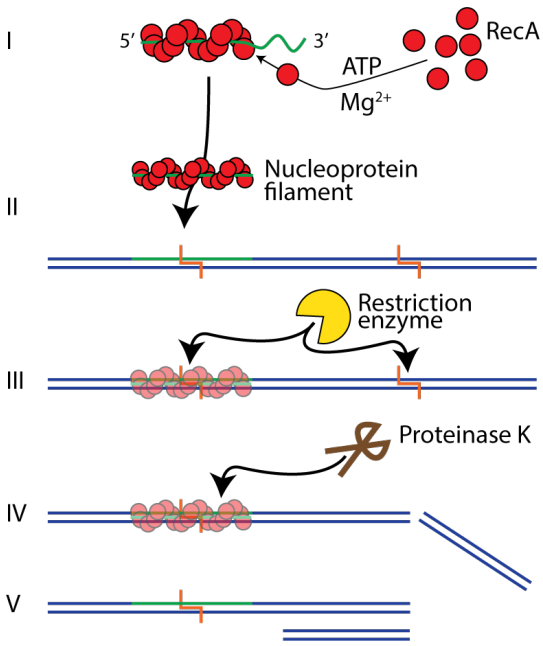

Figure 5.3: ***A schematic diagram depicting RecA mediated patterning and restriction assay. I)*** *RecA monomers polymerise on ssDNA in the presence of ATPγS and $Mg^{2+}$to form a NPF.* ***II)*** *The NPF is patterned across a restriction site preventing its cleavage* ***(III)****.* ***IV)*** *The NPF is removed using proteinase K for assessment by gel electrophoresis* ***(V)****.*

5.2.7 Nucleoprotein Filament Patterning Restriction Assay

In order to quantify the patterning efficiency an enzymatic assay was utilised, in accordance with Sharma et al.[312,6] Here, NPFs were patterned over unique restriction sites within the template dsDNA. The assembly of the NPF blocks the binding site, preventing the restriction enzyme from cutting the DNA template into shorter fragments (figure 5.3 III).

Removal of excess RecA; This process requires an additional step in the RecA patterning mechanism described above to ensure removal of excess monomeric RecA following the formation

of NPFs (section 5.2.6). Here, a 30 nt oligo containing only T's (d(T)30) is introduced at excess - typicaly 1 μl (@ 100 μM) - and the reaction mixture is further incubated for 15 minutes at 37$^{\circ}C$.

Restriction digest; Following synaptic joint formation (section 5.2.6) 2.5 μl restriction enzyme buffer (@ 10 x) and 1 μl restriction enzyme (@ 20 Units/μl) were added and the reaction volume made up to 25 μl with dH_20. The reaction mixture is incubated at 37$^{\circ}C$ for 30 minutes (figure 5.3 III). Note that the relevant restriction enzymes are specified in the text.

Removal of NPF; Finally all the protein - primarily the RecA - was removed from the reaction mixture prior to analysis via gel electrophoresis (section 5.2.4) (figure 5.3 IV). This was achieved through the addition of 5 μl of Proteinase K (@ 800 counts) accompanied by an incubation of 30 minutes at 37$^{\circ}C$.

5.2.8 Atomic Force Microscopy

In this thesis, the AFM was the primary interrogation tool and as such, a detailed examination of the working principles of the instrument was given previously in chapter 4. This section provides an overview of the sample preparation and operational parameters utilised in the work presented in the following chapters.

Sample Preparation

Sample preparation varied depending on the experiment. Broadly speaking there are three approaches utilised in this work; deposition for imaging in air; transient binding for imaging dynamics in liquid; and strong surface adsorption for high spatial resolutions in liquid. These are detailed individually below and variations from these protocols are specified where relevant in the text. Note that all sample incubations were conducted in a humidity controlled chamber.

Sample preparation for imaging in air; This approach is primarily utilised for the work demonstrated in chapter 6. Samples were appropriately diluted in either dH_2O or TE buffer (@ 1 x) to give a total deposition yield of 10 - 15 ng of DNA in a 20 μl volume. Muscovite Mica was freshly cleaved and incubated with 20 μl of 10 mM $NiCl_2$ for 2 minutes. Subsequently samples were deposited, incubated for 5 minutes, rinsed with a dH_2O (approximately 500 μl) and dried with N_2.

Transient surface binding for imaging in liquid; The development of this approach is detailed in chapter 7. Muscovite Mica was freshly cleaved and pre-incubated with $NiCl_2$ (20 μl) at

concentrations ranging from 1 mM to 15 mM for 1 minute. Excess $NiCl_2$ solution was removed and DNA solution (20 μl) was immediately applied to the mica surface and incubated for 5 minutes. Subsequently the surface was partially de-wetted. All DNA samples were prepared using 10 mM Tris-acetate (pH 7.4) and 10 mM Magnesium-acetate and 3.5 kbp DNA (10 ng). Mica samples were re-immersed in Tris-Mg buffer (200 μl) for imaging, with a small volume (50 μl) applied directly to the AFM cantilever. Where proteins or nucleoprotein complexes were introduced, they were appropriately diluted in Tris-Mg buffer and introduced directly into the imaging buffer column with a pipette, once stable imaging of the surface bound DNA substrate was established.

Strong surface binding for imaging in liquid; This approach is primarily used for the work detailed in chapter 9. Samples were diluted in 10 mM Tris-acetate (pH 7.4) and 20 mM $NiCl_2$ to give a total deposition yield of 10 ng of DNA - or equivalent NPF - in a 50 μl volume. Samples were deposited on freshly cleaved Mica and incubated for up to 90 minutes. Following this, the buffer volume was increased to 200 μl with 10 mM Tris-acetate (pH 7.5), to give a final Ni^{2+}concentration of 5 mM throughout imaging.

Where DNA origami was deposited, $Ni^{2+}$was found to cause detrimental effects when used for pre-incubation. $Mg^{2+}$was sufficient to enable equilibration between the tile and the surface to take place. 10 ng of DNA Origami tiles were deposited upon freshly cleaved Mica for 60 minutes in 10 mM Tris-acetate (pH 7.4) and 10 mM Mg-Acetate (20 μl). Where appropriate the introduction of 5 - 10 mM $Ni^{2+}$was found beneficial, post deposition.

AFM Operation

Operation in Air; All samples were imaged using either a Multimode Nanoscope IIIa AFM (Veeco, Santa Barbara, CA, USA) or a Dimension 3100, Nanoscope IV AFM (Veeco, Santa Barbara, CA, USA) in tapping mode in air. Images were acquired with an average scan rate of 1 Hz (512 x 512 pixels). All images were captured using OTESPA probes (Bruker, Santa Barbara, CA, USA) (Nominal values; L = 160 μm, W = 50 μm, K = 42 N/m, f_0 = 300 kHz in air, tip radius = 7 nm)

High speed AFM in aqueous buffer; All samples were imaged using a Dimension Fastscan (Bruker, Santa Barbara, CA, USA) in tapping mode in aqueous buffer. Images were acquired with typical scan rates between 38 and 60 Hz (256 x 256 pixels). All images were captured using Fastscan D probes (Bruker, Santa Barbara, CA, USA) (Nominal values; L = 16 μm, W = 4 μm, K

= 0.25 N/m, f_0 = 110 kHz in liquid, tip radius = 8 nm).

Rapid force-curve-based imaging in aqueous buffer; All samples were imaged using Peakforce tapping on a Dimension Fastscan (Bruker, Santa Barbara, CA, USA) using Fastscan D probes (Bruker, Santa Barbara, CA, USA) (values as above). The deflection sensitivity of each probe was calibrated using a 100 nm linear ramp on a mica substrate. Individual K values were established with a thermal calibration. Probes were engaged to the substrate with an engage setpoint of 0.1 V (approximately 1 nN) and a peakforce amplitude of 10 nm. Images were captured with a scan rate of 1 - 2 Hz (512 x 512 pixels) using a 5 nm peakforce amplitude, 2 kHz Peakforce frequency and a 1 μm Z range.

5.2.9 Design and Formation of DNA Origami

DNA origami structures were designed using the Cadnano software package.[314] The design process consists of; specifying the number of DNA duplexes across a square lattice (figure 5.4 A); routing the scaffold strand (figure 5.4 B); introducing staples (figure 5.4 C); populating the design with a DNA sequence (figure 5.4 D); and exporting the calculated staple sequences to be synthesised.

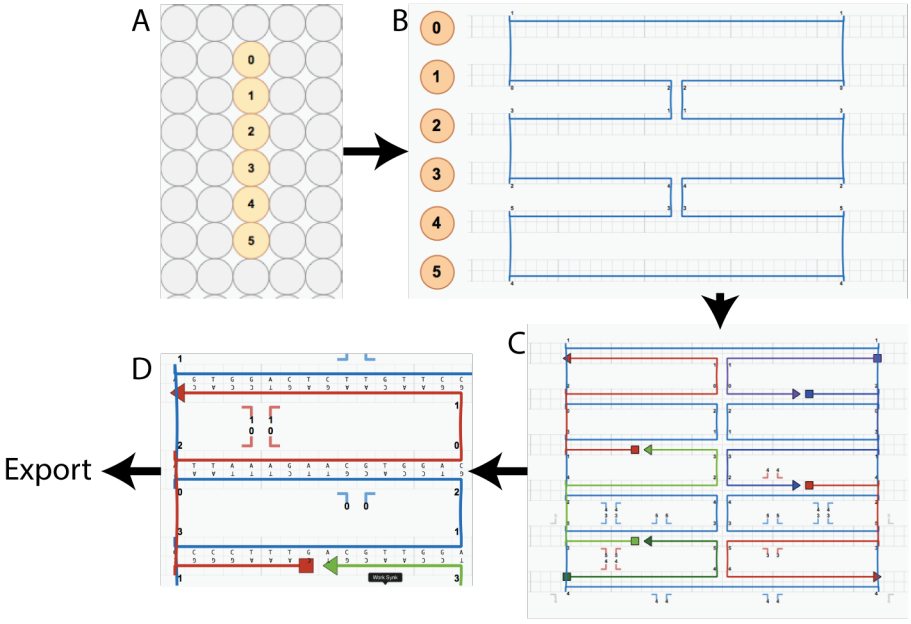

Figure 5.4: ***A overview of the design of DNA origami using cadnano A)*** *The number of dsDNA duplexes is specified on a square lattice.* ***B)*** *The scaffold (blue) is rastered through each duplex to create the two dimensional shape.* ***C)*** *The design is populated with staple strands (various colours).* ***D)*** *Finally, the nucleotide sequence of the scaffold strand is populated, from which the staple sequences are derived.*

Folding

The origami structures were generated using M13mp18 ssDNA (New England Biolabs, MA, USA) scaffold with an average of 220 staple oligonucleotides (IDT, IA, USA), each typically 32 nt long. Staple mixtures were produced to a final concentration of 100 nM (final concentration of each oligonucleotide) in 10 mM Tris-HCl or Tris-Acetate (pH 7.4), 1 mM EDTA and 10 mM $MgCl_2$ or Mg-Ac. A typical folding reaction consisted of 2 μl M13mp18 ssDNA (@ 100 nM), 10 μl Staple mix (@ 100 nM), 1 μl Tris-Ac (pH 7.4) (@ 200 mM), 2 μl EDTA (@ 10 mM) and 2 μl Mg-Ac (100 mM) made up to 20 μl with dH_2O. Note, DNA origami was also found to fold perfectly without EDTA. Reactions were heated to 95$^\circ C$ for 5 minutes and cooled using a controlled gradient of -1$^\circ C$ min^{-1} to room temperature.

Inclusion of Central Strand

Regarding the DNA origami frame structures, the central dsDNA strands were incorporated in excess following initial folding of the frame. Incorporation was usually carried out at between 2 - 5 : 1 (central strand : frame) ratios. A typical reaction consisted of; 20 μl DNA origami frame (@ 10 nM), 2 μl dsDNA central strand(s) (@ 200 - 500 nM), 1 μl Tris-Ac (@ 200 mM), 2 μl EDTA (@ 10 mM) and 2 μl Mg-Ac (@ 100 mM) made up to 40 μl with dH_2O. Reactions were heated to 40 - 50$^\circ C$ - depending on the T_m of the anchor sequences - and cooled using a controlled gradient of -0.5$^\circ C$ min^{-1} to room temperature.

Purification

All origami structures were purified using Illustra MicroSpin columns (GE Healthcare, UK) containing a Sephacryl size exclusion chromatography matrix (figure 5.5). Note that, different size exclusion limits were required for the different origami frames, typically Sepahcryl S-400 was utilised. This technique separates molecules based on their Molecular weight (MW). Larger molecules - in this case the DNA origami - pass through the column in the excluded volume (figure 5.5, green). In comparison, small molecules are able to enter the pores of the Sephacryl matrix and thus have a larger interaction volume to diffuse through. These molecules are fractioned based on their relative MW where the largest elute first (figure 5.5, yellow and red).

Prior to use, the storage buffer was removed from the matrix by centrifugation for 1 minute. Each column was equilibrated with 200 μl Tris-Mg buffer three times prior to the introduction of any

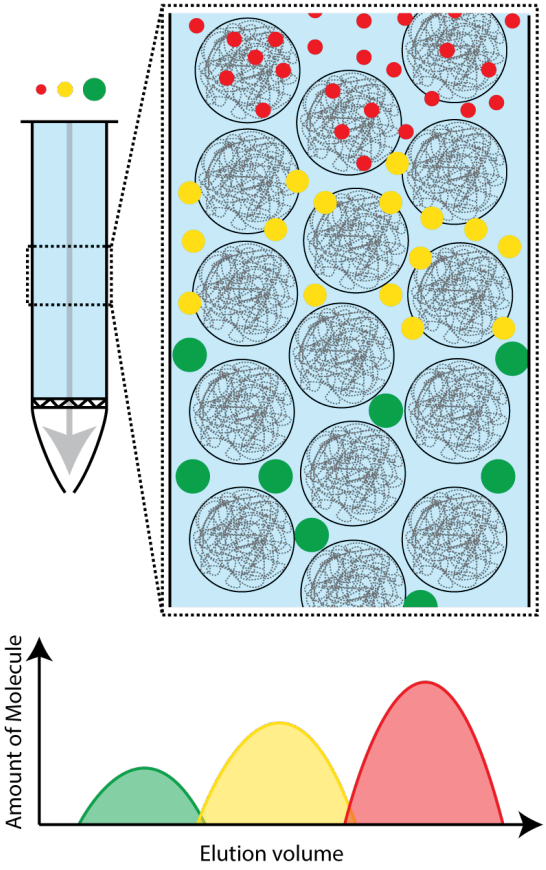

Figure 5.5: ***A schematic diagram depicting size exclusion purification columns.*** *The column is packed with a sephacryl based porous gel matrix. A population of molecules of different molecular weights is separated due to their differential access to the pore volume of the matrix. Molecules too large to fit into the pores (green) are eluted first in the exclusion volume. The remaining molecules are fractioned based on their dwell time within the matrix volume. A representative elution profile is given.*

sample. This buffer was applied to the top of the column - with care taken not to disrupt the matrix - and centrifuged for 1 minute, with the elutant being discarded. Samples were applied to matrix in the same manner and the elutant collected in a 1.7 ml eppendorf tube following centrifugation for 30 seconds - 1 minute. All spin steps were carried out in a tabletop microcentrifuge (eppendorf) at 750 rcf (relative centrifugal force) at room temperature. Typically samples were pooled for purification and 20 μl aliquots were frozen for later use.

This page is intentionally left blank.

Chapter 6

Examining the Interaction Landscape of RecA Nucleoprotein Filaments

6.1 Introduction

As highlighted in chapter 3, the mechanism by which RecA NPFs search for homologous sequence amongst vast heterology continues to remain unclear. This discrepancy in understanding is a consequence of the indirect single molecule investigation methodologies employed to date, which are insensitive to the presence of any intermediate states (section 3.4.2). So far, three states have been characterised from averaged structural studies, including; pre-synaptic, synaptic and post-synaptic complexes (section 3.2.3).

With respect to the homology searching mechanism, the synaptic state is proposed to represent a transient interaction between heterologous sequence pairings. This integral state is characterised as the central "homology search" process which transitions to a post-synaptic state upon full homologous sequence pairing (section 3.2.3).

In this chapter, the prevalence of the synaptic and post-synaptic intermediate states is statistically assessed across a typical patterning reaction window of 60 minutes (see methods 5.2.6), which is temporally dissected into "snap shots". Such snap shots are crucial in reconciling the constituent intermediate states that characterise the RecA homology searching mechanism and will inform any later dynamic observations made with HS-AFM.

As described previously (section 3.3.1), the Bioelectronics group has been developing methodologies to harness the RecA mediated homologous recombination mechanism in order to

spatially address synthetic DNA structures for bottom-up assembly of nano-devices and functional materials.

As such, there are several well established protocols within the group for the formation of RecA NPFs and the patterning of dsDNA templates. To date, the main assessment of which has been carried out utilising a restriction enzyme based biochemical assay (see methods 5.2.7), enabling quantitative analysis of terminal patterning efficiencies.[312,6] These studies, have been further complimented with AFM observations in air to confirm the terminal reaction state.

These approaches rely on solution based interactions with high component concentrations and stoichiometries to enable the high efficiencies described to date.[312,6] Although perfectly suited to the intended construction approach, these reaction conditions pose severe complications – in particular surface fouling – where direct single molecule observations are intended, such as in the latter parts of this thesis.

In contrast, in this chapter low stoichiometric NPF:dsDNA interactions are first characterised with the established restriction assay (see methods 5.2.7). These conditions are subsequently applied to reaction snap shots – from 5 - 60 minutes – assessed with the AFM in air. Furthermore, in line with the intended HS-AFM observations – and subsequent bionanotechnological applications – the affect of temperature, NPF size and the presence of a solid-liquid interface on the reaction landscape are all investigated. These experiments aim to establish a base set of interaction intermediates from which all subsequent HS-AFM observations can be compared – taking into account system constrains, such as the presence of a solid-liquid interface.

6.2 Experimental

6.2.1 DNA Substrate Design and Generation

In order to investigate the interaction landscape a simple linear dsDNA template was required, sufficiently long to be easily and unambiguously identifiable by AFM and yet short enough to be well resolved in a gel electrophoresis setup. Previous work by others in the Bioelectronics group had utilised a 890 bp template for these purposes, which was deemed suitable for this preliminary work. This dsDNA template was easily identifiable by AFM at 302 nm in length, and contained a unique BanI restriction site 350 bp from the 5' termini of the leading strand. The asymmetrical design was argued to enable less ambiguous distinction between the synaptic and the postsynaptic joints as the site of homology was kept 85 bp away from the center (shown in figure 6.1).

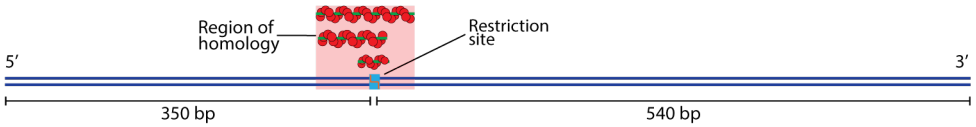

Figure 6.1: ***A schematic diagram depicting the layout of the 890 bp template.*** *The locations of the BanI restriction site and relative fragment sizes is indicated. The region of homology covers the restriction site, protecting it from enzyme access. The relative positions of the 90, 60 and 30 nt homologous nucleoprotein filaments is shown, top - bottom, respectively.*

A set of 30, 60 and 90 nt ssDNA oligonucleotides were generated to pattern the 890 bp template, sharing sequence homology with the upper strand. These regions of homology were designed to pattern across the denoted restriction site, thus protecting it from enzymatic cleavage, in accordance with the protocol described in methods 5.2.7. NPFs formed from ssDNA of this length have been shown to be highly efficient at forming synaptic joints[6,315] and remain consistent with those used in previous FRET based studies.[16] Furthermore, at 15 nm in length and 10 nm in width, 30 nt NPFs were considered the smallest feature that could be resolved unambiguously in the 5μm x 5μm (512 x 512 pixels) images used in these experiments.

6.2.2 Assessing the RecA Nucleoprotein Filament Homology Search by Enzymatic Assay

First and foremost, all components were tested to ensure correct patterning at the low stoichiometries required by the AFM experiments that follow. In accordance with previous work by Sharma et al[4,5,6] high patterning efficiencies are achieved using a typical stoichiometry of 40:1 (NPF:dsDNA), with excess RecA bound to an additional heterologous 30 nt oligonucleotide prior to the introduction of a dsDNA substrate. However, this approach would not be appropriate for single molecule observations, leading to large background contamination. As such, the patterning efficiency of a more practical 3:1, stoichiometry was explored.

NPFs were formed on the 30, 60 and 90 nt oligonucleotides in the presence of ATPγS and 2 mM Mg^{2+}in accordance with methods 5.2.7. The resultant NPFs were subsequently introduced to 890 bp dsDNA at a 3:1 ratio and incubated for 1 hour (methods 5.2.6). Subsequently, all DNA species were digested with BanI for 30 minutes and followed by complete protein removal by Proteinase K incubation (methods 5.2.7).

From figure 6.2, clear protection of the 890 bp template is observed as a result of the NPFs presence, with efficiency varying with respect to length. Where various components of the NPF

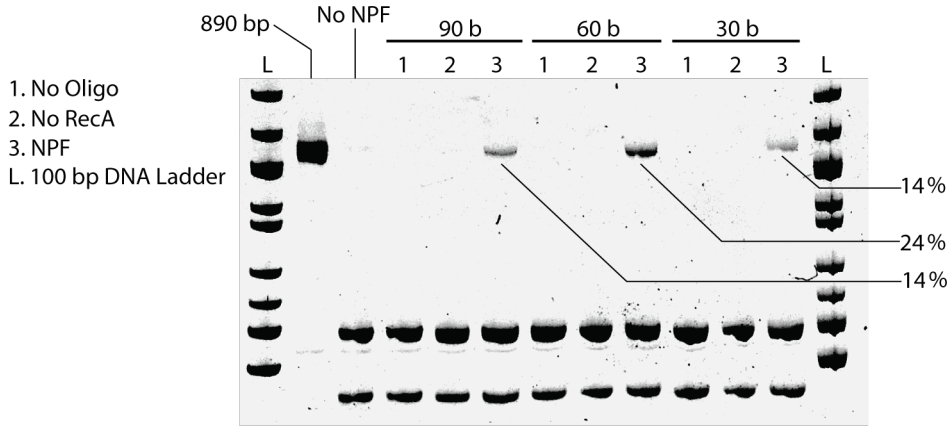

Figure 6.2: ***RecA restriction digest gel of 890 bp.*** *A 5% polyacrylamide gel depicting the restriction assay of 890 bp patterned with 30, 60 and 90 b NPFs at a 3 to 1 ratio. Respective estimated patterning efficiencies are given as densitometry analysis of the relevant lanes. L = 1 kbp ladder.*

are omitted (figure 6.2 lanes references: No NPF, 1 & 2) then no protection from the restriction enzyme is afforded and the 890 bp DNA is cleaved into two fragments; 540 bp and 350 bp (figure 6.2).

Interestingly, over the course of 60 minutes the 60 nt NPF patterns with 24 % efficiency at the 3:1 ratio used here, compared to 95% at a 40:1 ratio.[312,4] This yield is reasoned to be sufficient for identifying interaction intermediates at the single molecule level, with a quarter of all observed molecules under the AFM expected to display a successful interaction. In contrast, the 30 nt NPF is found to yield a reduced efficiency of only 14 %. This difference from that of the 60 nt is notable, where one may postulate that patterning efficiency decreases with NPF size - indeed this was the case noted in the earlier work of Sharma et al,[6] where filaments were scaled down to 6 nt. However, this argument is partially undermined by the apparent similar efficiency of the 90 nt NPF. Nonetheless, all patterning efficiencies were deemed to be of sufficient yield for subsequent investigations with the AFM.

6.2.3 Observing the RecA Nucleoprotein Filament Intermediates by AFM

In order to interrogate the nature of intermediates formed during the search for homologous sequences, a 60 minute interaction time course was temporally dissected to statistically sample the reaction intermediates. 60 nt NPFs were formed as previous and incubated with 890 bp dsDNA at $37°C$ (methods 5.2.6).

The reactions were terminated by quenching the reaction mixture to $4°C$ at regular time intervals, starting from 5 minutes and going up to 60 minutes. This temperature was found to be sufficient for ceasing any interactions between the NPFs and the dsDNA scaffold (Figure 6.3 B), with only 6.4 % of the surface bound DNA population indicated the formation of a synaptic – and no post-synaptic – complexes at 60 minutes compared to approximately 40 % at $37°C$. From figure 6.3 B it can be seen that the majority of the NPFs are present within the background when incubated at $4°C$, indicating limited - if any - activity. Given this, samples were quenched and diluted by the addition of dH_2O at $4°C$.

Typically, a total of 10 ng of dsDNA was deposited upon Ni^{2+}pre-incubated mica for 5 minutes, prior to rinsing with dH_2O and drying with N_2 (section 5.2.8). Following this samples were visualised with AFM, using tapping mode in air (methods 5.2.8) - for any intermediates formed between the NPF and dsDNA.

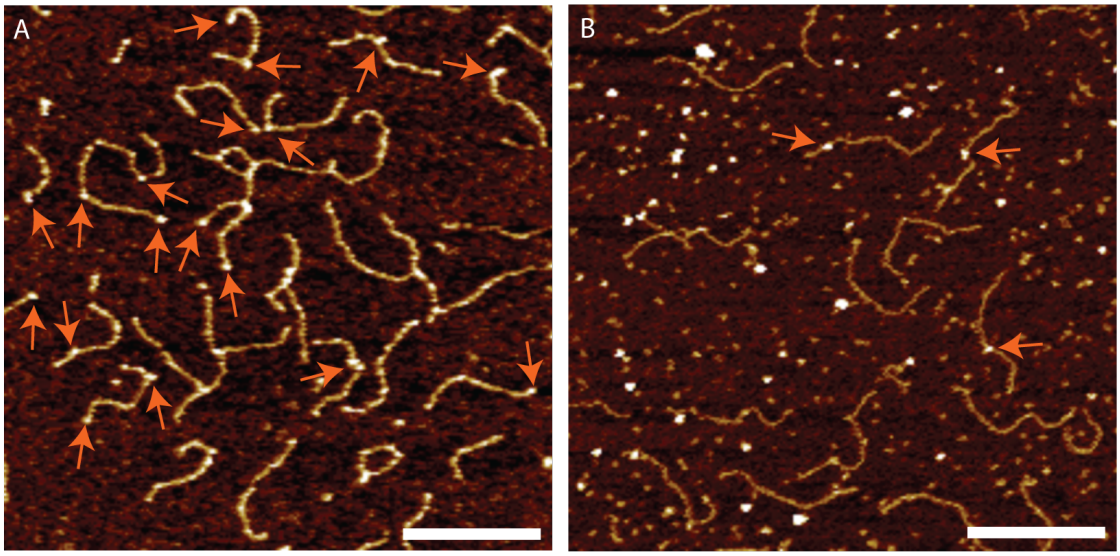

Figure 6.3: ***A comparison of patterning efficiencies at 37°C and 4°C .*** *60 nt NPF (orange arrows) are shown to interact with 890 bp at 37°C* ***(A)*** *but to an extremely limited degree at 4°C* ***(B)****, where the majority of NPFs are observed in the background. Scale bars = 250 nm. Z scale = 8 nm.*

Substrates were imaged at random locations, in order to ensure a large and unbiased population of DNA molecules were sampled, typically n = 1000 DNA molecules per sample. Only DNA molecules who's termini could be unambiguously discerned, that fell entirely within the imaging area and that did not overlap with one another were included in the population count. Great care was taken to keep consistent concentrations of each molecule throughout the reaction time series to ensure repeatable mono-dispersions of reaction intermediates for microscopic observation.

The resultant population of dsDNA observed were categorized as a function of the observed interaction state of the NPF with the 890 bp DNA, i.e. those which; a) did not interact with the DNA; b) perhaps formed a synaptic joint with the scaffold; c) had already resolved into postsynaptic complex at the correct homologous site – based upon position on the asymmetric DNA scaffold.

A representative set of AFM images is shown in figure 6.4 B - E, where in panel B a successful postsynaptic complex can be seen on 890 bp DNA. Other panels, (C - E) represents 890 bp DNA with multiple NPFs associated at heterologous locations.

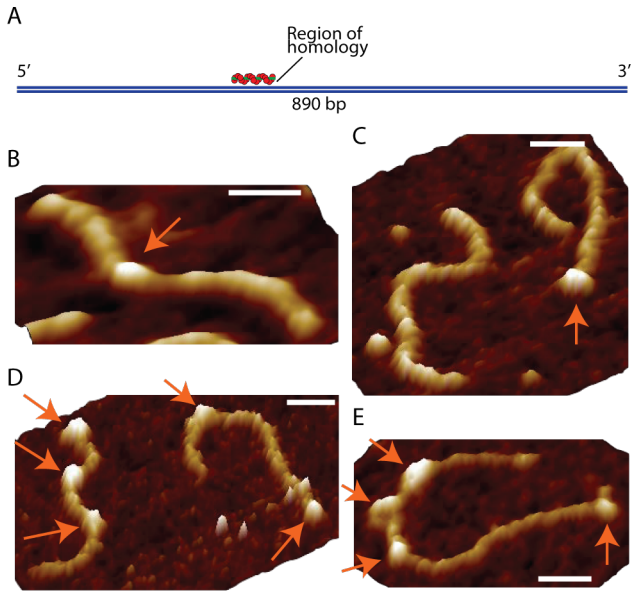

Figure 6.4: ***A montage of AFM images depicting parallel RecA homology searching.*** ***A)*** *a schematic diagram depicting the region of 60 nt homology on a 890 bp template.* ***B - E*** *shows the variation of reaction intermediates, with;* ***B)*** *a successfully formed synaptic joint at the point of sequence homology;* ***C)*** *free filaments and DNA alongside terminally bound DNA template; and* ***D) & E)*** *demonstrate multiple pre-synaptic joints formed along their length. Nucleoprotein filaments are highlighted with orange arrows. All Scale bars = 60 nm. All Z scales = 8 nm.*

From the observations made here, it can be seen that NPFs interact with dsDNA in a largely parallel process. This is in stark contrast to previous studies which are insensitive to these additional and intermediate interactions (see section 3.4).

The parallel nature of the homology searching mechanism results in a multi-phase mechanism, which is evidenced across the full 60 minute time course. From figure 6.5 A, two phases are clearly observed: a rapid first phase characterised by the parallel formation of synaptic joints – referred to here as the 'association' phase; and a slower second phase characterised by

the dissociation of weakly bound heterologous synaptic joints and an increasing population of postsynaptic complexes – referred to as the 'resolution' phase.

Rapid Association Phase

For a reaction carried out at 37 °C, it was found that within 10 minutes the percentage of the 890 bp DNA population displaying an association of some nature with a 60 nt NPF rapidly saturated to $91.6 \pm 1\%$. Of this only a small population $4.4 \pm 1\%$ had postsynaptic joints formed at the correct homologous location, while the remaining $87.2 \pm 1\%$ had one or more synaptic joints formed at heterologous locations. This is shown in the histogram of figure 6.5 A, green bars. Interestingly the vast majority of DNA molecules were observed to have multiple interactions simultaneously (figure 6.4). On average, 3 synaptic joints per 890 bp DNA were observed, with up to 7 seen on a few of the templates. Interestingly, the latter equates to approximately half the available template base pairs being probed concurrently.

From these observations it becomes immediately apparent that the homology search begins in a largely parallel fashion. In contrast, previous studies are only able to report findings based on the interactions between a single NPF and dsDNA due to indirect reporting methods (see section 3.4). It is not hard to see how the multiplexing of the homology search effort would increase the efficiency of the process, especially across a large sequence space such as a genomic DNA. Indeed, it could be argued that it is a logical component of any searching mechanism that is entirely thermally driven. Whilst this mechanism does not save on energy expenditure, it would decrease the searching time by a factor of 7, for example, where seven synaptic joints are formed concurrently and allowed to undertake short regions of sliding. As expected the chances of filaments to find the region of homology at the first encounter is extremely low even for short templates such as the 890 bp DNA used here.

A similar saturation profile was observed by 20 minutes for samples incubated at $22^{\circ}C$. Of these, $54.5 \pm 3\%$ of the population demonstrated a reaction intermediate, of which only $3.6 \pm 1\%$ was found to be a postsynaptic joint (Figure 6.5 A, blue). This further supports the previous suggestion of a thermally driven searching mechanism, although whether this is as a direct consequence of reduced diffusion of either the NPF, the dsDNA into the complex or a reduction in local DNA breathing – partial separation of the DNA strands due to thermal fluctuations – is unclear, perhaps a combination is likely.[173] The nature of parallel interactions observed within this system raises some interesting questions. What happens when one of the multiple synaptic joints finds

a region of homology? How do the other synaptic joints resolve and is this even necessary? Can unresolved joints lead to irregular and detrimental recombination events between mostly heterologous partners?

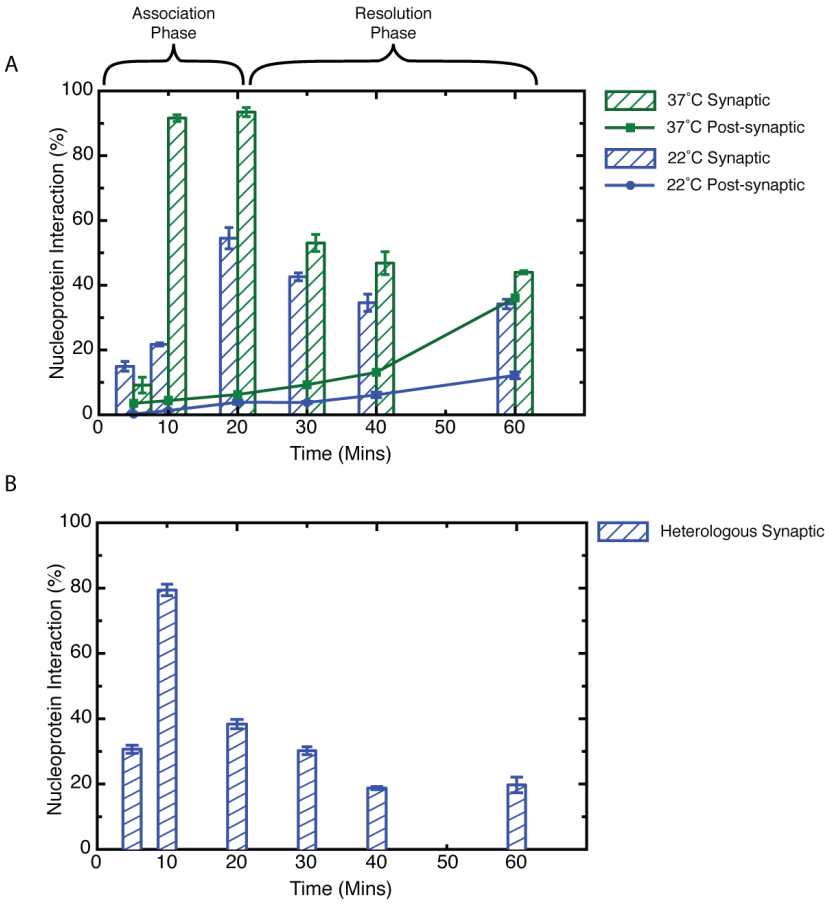

Figure 6.5: ***The multi phase interaction landscape of RecA NPF homologous recombination.*** *A set of histogram plots comparing the interaction landscapes of 60 nt homologous* ***(A)*** *and heterologous* ***(B)*** *NPFs.* ***A)*** *depicts the bi-phasic reaction landscape at 37 °C (green bars) and 22 °C (blue bars) revealing a largely parallel pre-synaptic joint formation phase and a synaptic joint resolution phase (green & blue lines) with a strong dependence upon the thermodynamics of the environment.*

Resolution Phase

Interestingly, following the saturation of dsDNA-NPF intermediates a second phase becomes evident from 10 to 60 minutes in which the population of dsDNA demonstrating a reaction intermediate rapidly decreased to 43% ± 0.51% (Figure 6.5 A). Simultaneously, this is associated with an increase in the percentage of the postsynaptic population, which steadily increases across 60 minutes reaching 36% ± 0.58%. This is attributed to formation of a stable three-stranded

postsynaptic joint when one of the synaptic intermediates successfully locates the homologous site, while the relatively weak heterologous synapses dissociate from the DNA duplex.

Similar behaviour was observed for a heterologous 60 nt NPFs, where roughly the same number of synaptic joints was seen to form during the early stages of interaction. Further, these were found to drop off drastically during the resolution phase, without the associated postsynaptic transitions (Figure 6.5 B). Moreover, when samples were imaged after an overnight incubation, only bare 890 bp DNA were observed (figure 6.6) due to the disassembly of the NPFs following ATPγS hydrolysis – which proceeds at 0.01 min^{-1} .[316]

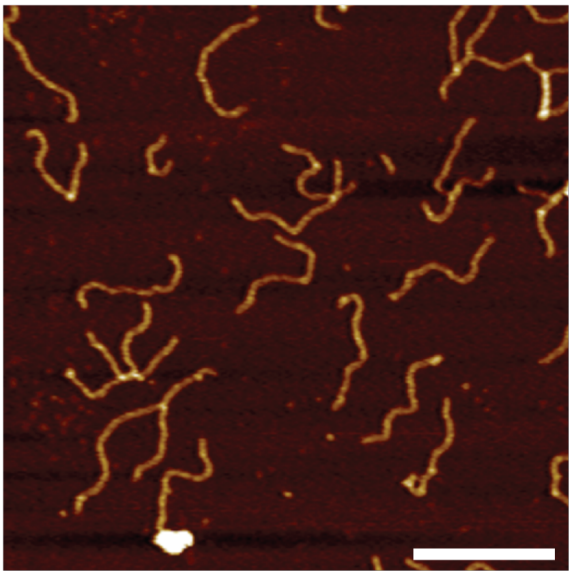

Figure 6.6: ***Disassembly of 60 nt NPF patterned on 890 bp following incubation at 37°C for 24 hours.*** *Scale bar = 250 nm. Z scale = 8 nm.*

It has previously been proposed that the NPF construct interacts non-specifically with the dsDNA template through the exterior of the filament and that it is the DNA which wraps in to the helical groove of the complex - where the ssDNA nucleotides are exposed - through thermal fluctuations alone.[160] If a partial match is achieved the complex becomes stable and facilitates further base probing, progressing at a theoretically exponential rate. However, if no homology is present the complex remains unstable, with the dsDNA free to unwrap from the NPFs helical groove.[161,221] The mobility of the dsDNA to come into and out of the helical groove of the NPF is therefore dependent upon the thermodynamics of the system requiring more energy to leave the complex than to enter it, which explains the formation and rapid disassembly of transiently bound synaptic joints during the early phase of homology search in these results.

Hence, it can be postulated that in the beginning of the homology search, irrespective of the

sequence, the nucleoprotein filament docks on the dsDNA to form a synaptic joint. Association is however short lived unless homology - at least in part - is located and a full transition to a post-synaptic joint occurs.

6.2.4 Nucleoprotein Filament Size and Homology Search

Previously as few as 6-8 nt short filament were found to confer stable postsynaptic joints, but with reduced efficiency when compared to longer filaments.[6,194,195] It could be argued that this is related to the increased stabilisation offered by the larger surface areas of longer filaments through which a greater number of non-specific interactions with the duplex DNA can occur. Having said that, this does not necessarily translate to an increase in postsynaptic joint formation efficiency, where an increase in transient joint stability could be counteracted by an increase in the MW of the NPF, and hence a reduction in the rate of diffusion. Hence, in order to further investigate the effect of the filament length on the nature of intermediates formed during the homology search, a set of experiments utilising 30 nt and 90 nt long NPFs homologous to the same region of 890 bp DNA were conducted. The results are shown in Figure 6.7.

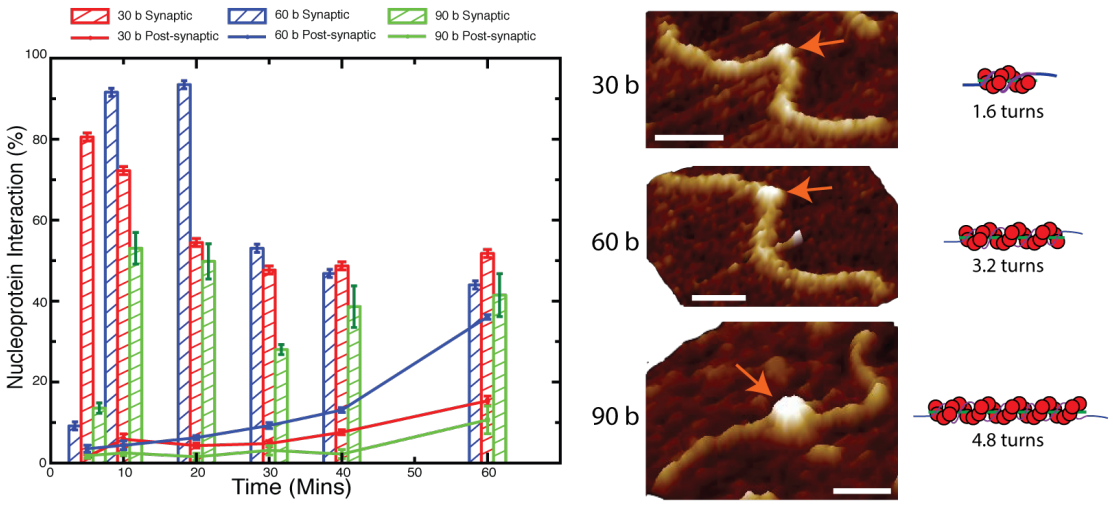

Figure 6.7: ***The size dependence of NPF homology searching.*** *30 (red), 60 (blue) and 90 nt (green) NPF demonstrate differential rates of pre-synaptic and synaptic joint formation. This is attributed to two factors, increase in molecular mass resulting in a decrease in diffusion and a reduction in the pre-synaptic joint stability as the NPF complex size diminishes. These effects combined result in a greater dynamic turnover of pre-synaptic complexes with reduced filament size. Larger filaments are theorised to be stable for longer on dsDNA through non-specific contact. Scale bars = 60 nm. Z scales = 8 nm.*

During the association phase, the number of synaptic joints formed by the shorter 30 nt NPF was

seen to peak even after 5 minutes of incubation (80.6 $\pm$ 2%), compared to 10 – 20 minutes seen for 60 nt and 90 nt filaments (Figure 6.7). This was associated with a very rapid second phase, resolving to approximately consistent 50% of the DNA population representing a nucleoprotein interaction after 20 minutes (Figure 6.7, red bars). A similar plateau was observed after 40 minutes for 60 nt and 90 nt NPFs, suggesting that synaptic joints formed with longer filaments were far more stable.

It is important to note that for 30 nt ssDNA only 10 RecA monomers are needed to form a NPF which consists of only 1.6 helical turns; compared to 20 and 30 RecA monomers required for 60 and 90 nt long ssDNA respectively. This is shown in Figure 6.7, along with the 3.3 and 4.9 helical turns of a 60 and 90 nt NPFs, respectively. Hence, it is not surprising that a 30 nt NPF is considerably less stable than its 60 nt or 90 nt equivalent, therefore, able to form multiple nascent synaptic joints with very short lifespans. This is in-line with earlier reports which suggest an increase in the lifespan of synaptic intermediates for either long filaments.[15,17] As a consequence short NPFs are able to undergo a more rapid resolution phase.

When the percentage of 890 bp DNA with postsynaptic complexes from each population is considered, a remarkable reduction in the yield of successful postsynaptic joints was observed. For the 30 nt NPFs, only 15.4 $\pm$ 1.2% of the population had postsynaptic joints formed at the correct location, found to be in line with the efficiency reported by the previous restriction assay (section 6.2.2). This represents less than half the efficiency of the 60 nt NPFs at 36% $\pm$ 0.58% (Figure 6.7). Further, the efficiency of 90 nt NPF postsynaptic joint formation after an hour was found to be only 10.66% $\pm$ 3.6%. Interestingly, this was similar to the yields reported for 60 nt NPFs at 22$^{\circ}C$ (Figure 6.5 A, blue). This fits well, given that a 90 nt NPF represents 1.5 times the MW of a 60 nt NPF, compared to the a 22$^{\circ}C$ system which represents $\frac{2}{3}$ the thermal energy of the optimal interaction energy (at 37$^{\circ}C$). Interestingly it is noted that, despite a large difference in the over all interaction intermediates between 30 and 90 nt NPFs, a similar proportion of postsynaptic joints were transitioned by 60 minutes (Figure 6.7).

This observation at first maybe misconstrued as resulting from the same effect. However, it has previously been demonstrated that there are no conformational changes or turnover of chemical energy products within the homology searching mechanism.[317,204] One must therefore consider that unsuccessful synaptic complexes are not spent, but may be able to undergo further interactions immediately. Therefore, a rapid turnover of shorter, less stable NPFs would suggest a system where for the same thermal energy more interactions can be achieved in the same temporal frame. Thus, the interacting population resolves to a roughly steady state population of synaptic joints,

with a small proportion of successful postsynaptic complexes represented in the case of 30 nt NPFs. In contrast, 90 nt NPFs represent the opposing end of this process, where an increase in the size negatively impacts the overall efficiency, due to the decreased diffusion and subsequent turnover rates. The net result of these processes can be considered to give broadly the same results, as observed in figure 6.7.

6.2.5 Homology Searching in the Context of Persistent Confinement at a Solid - Liquid Interface

With regards to exploring the underlying biological mechanism, one must consider the RecA mediated homology search in the context of a crowded cellular environment, where the persistent confinement of genomic DNA exists in tandem with a plethora of other nucleoprotein interactions (Figure 6.8 A).[318]

Interestingly, knowledge of reactions states under such consistent confinement is also critical when attempting to orchestrate RecA mediated patterning within complex DNA topologies, typically at a solid–liquid interface or during dynamic studies with the HS-AFM (figure 6.8 B).

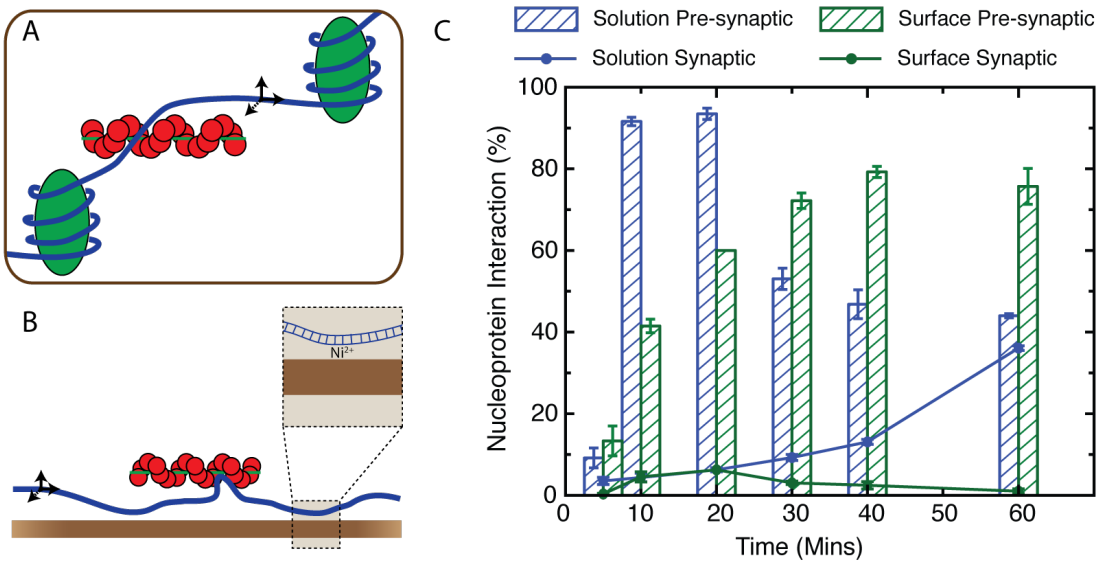

Figure 6.8: ***NPF homology searching within the context of persistent template confinement. A)*** *homology probing within the context of a cellular environment and B) a mimicked confinement regime. The histogram in* ***(C)*** *depicts a break down in the bi-phasic reaction landscape as a result of stabilised pre-synaptic filaments in a confined regime (green).*

To this end, 60 nt filaments were introduced to 890 bp DNA which had been pre-immobilised on a mica surface - pre-incubated with 10 mM Ni^{2+}- and allowed to interact at $37°C$. Under these

conditions the dsDNA is bound to mica at various points through association with the Ni^{2+}cations (figure 6.8 B), this is examined in detail in the next chapter .[319,320] In this way, the surface confined DNA to a certain extent is considered to mimic the structural confinement of genomic DNA or complex DNA topologies (figure 6.8 A). These samples were observed, as previous, with the AFM in air following drying at the designated time point. The results are shown in Figure 6.8 C.

RecA NPFs were found to interact successfully with partially immobilised DNA at a solid-liquid interface. Of these surface bound populations, the saturation of synaptic joints at 79.2% $\pm$ 1.4% was observed only after 40 minutes (figure 6.8 C, green bars). Unsurprisingly, this represents a reduction in the interaction rate compared to the identical case in a solution-based system, suggesting that DNA accessibility is limiting, in-line with previous discussions (section 3.2.3). However, one must also consider that NPFs may become immobilised upon the surface themselves, hence lowering the interaction stoichiometry.

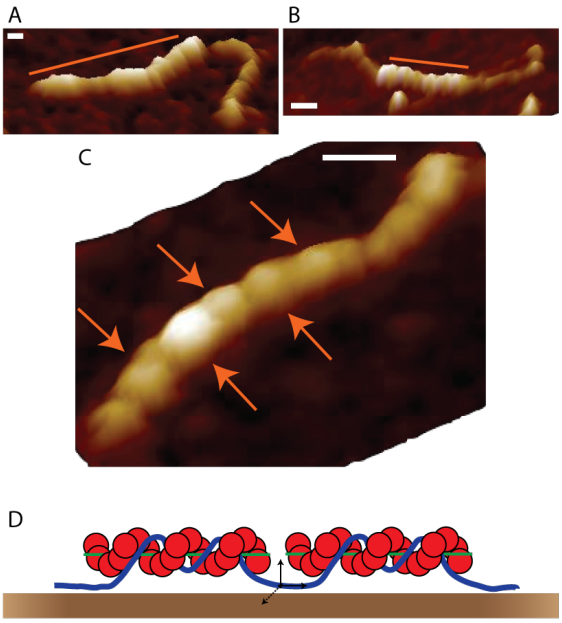

Figure 6.9: ***The observation of bunched RecA NPFs on 890 bp.A & B)*** *A set of AFM images depicting characteristic large complexes (orange lines) that dominant reaction intermediates within a confinement regime. These are revealed to be "bunches" of 60 nt NPFs* ***(E, orange arrows). F)*** *Proposed 60 nt NPF "bunches" shown schematically. Scale bars = 60 nm. Z scales = 8 nm.*

From Figure 6.8 C, a single-phase saturation event is evident, in contrast to the multi-phasic trends seen in solution-based experiments. This demonstrates a departure from the resolution phase suggesting a stabilisation of the synaptic joints formed on dsDNA due to the spatial confinement of the solid-liquid interface. As a result the synaptic joints were observed to be "trapped" on

surface-bound DNA templates, as the individual NPFs are unable to dissociate. This can clearly be seen from the AFM images in figure 6.9 A & B (orange lines), where characteristic "bunches" of synaptic joints are observed on 890 bp DNA. Furthermore, individual 60 nt NPF can be easily resolved in these bunched complexes (Figure 6.9 C, orange arrows). Hence, this trapping of NPFs on the DNA template resulted in the decline of observed postsynaptic joints to approximately 1% of the population; despite an early increase (Figure 6.8, green line).

It is suggested that these "bunches" NPF complexes represent the stabilisation of the otherwise weak synaptic joints, which are unable to resolve as a consequence. It is postulated that the surface confined dsDNA is further stabilised by the non-specific interactions within the helical groove of the bound NPFs. As a consequence of the reduced degrees of freedom, the DNA is unable to diffuse out to unwrap from the NPF. This situation, is expected to be further compounded as the number of NPFs associated increases, further stabilising the DNA within the growing complex (Figure 6.9 D). These direct single molecule observations are in support of the previous work suggesting that for RecA nucleoprotein filaments to interact with DNA, the three dimensional degrees of freedom of the latter molecule are of critical importance.[14]

These observations further demonstrate the parallel nature of the RecA NPF homology searching interaction, such that multiple synaptic complexes are simultaneously found to be probing the template DNA sequence. Importantly, previous work has suggested that the homology searching rate is independent of template length or topology,[321] a point which is corroborated by this work given the scalable nature of a parallel searching mechanism.

6.2.6 RecA as a Pack Hunter

Surprisingly, throughout the above experiments significant evidence for the existence of NPF occurring in clusters has been noted. These observations are intriguing, as the appearance of tightly clustered NPFs – seen to interact with the 890 bp DNA templates – can be directly reconciled with the parallel multi-phase interaction profiles described in this chapter. This can clearly be seen in figure 6.10 for 60 nt NPFs where the majority of NPFs are observed to exist in highly concentrated arrangements, suggesting that RecA has some distinct association behaviour. These clusters of NPFs are contrary to the expected random dispersion of NPFs, and are prominently distinguishable from the RecA background.

When taken together with the work presented in the previous sections, it is evident that the observed parallel nature of the RecA homology searching mechanism is orchestrated - at least

in part - by the occurrence of these cluster formations. One may consider that these clusters act to locally enhance the NPF concentration in the vicinity of the DNA, facilitating simultaneous NPF interactions. Interestingly, NPFs have previously been shown by others to associate with one and other - in some cases, forming bundles and aggregates.[322] Furthermore, these bundles have been shown to form *in vivo* in bacteria where it is suggested that they accelerate homology pairing between distant loci.[323] These studies, therefore, would largely support that the clusters observed in this work arise from a similar phenomena.

Still, a key difference in the results presented here, is that the NPFs are observed to be loosely associated, appearing to be cooperatively interacting within the same local focal point. This is in contrast to previous work, which suggest higher order bundle structures arise from direct contact between neighbouring NPFs - primarily through the protruding C terminal (section 3.2.2). Hence, the two processes must be considered separate and may have unique roles in the process of homologous recombination.

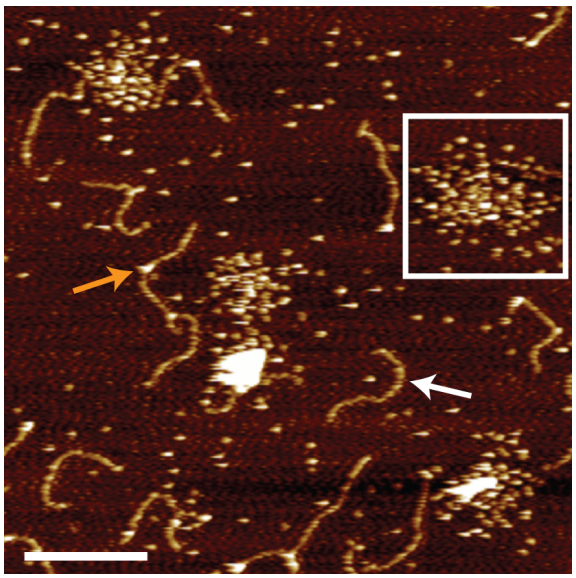

Figure 6.10: ***A Parallel homology search orchestrated by cooperative NPF clustering.*** *A representative AFM image depicting the observation of 60 nt NPF clusters gathered in the vicinity of 890 bp dsDNA. Scale bars = 125 nm. Z scales = 8 nm.*

6.3 Conclusions

The work in this chapter has directly interrogated the RecA homology searching mechanism at a single molecule level, using the AFM to statistically assess the nature of intermediates formed throughout a 60-minute reaction window.

From this work it has been revealed that the interaction landscape is considerably more complex than demonstrated to date, where previous methodologies were limited either by the spatial resolution or the flexibility to distinguish multiple events. In this chapter it has been shown that the RecA mediated homology search occurs in multiple phases, necessitated by the massively parallel interaction regime.

As such, NPFs are shown to interact with a dsDNA in a cooperative manner resulting in a rapid 'association' phase where multiple synaptic joints are formed per DNA molecule. This association, is however, short-lived if the complexes are not stabilised by the location of homologous sequence. This is most profound for shorter NPFs. At this point the second 'resolution' phase becomes prevalent as synaptic joints, transiently bound at regions of heterology, dissociate.

The number of postsynaptic complexes was typically found to increase over the course of the reaction, as heterologous joints were resolved (see figure 6.5). Finally, all complexes - including the postsynaptic joints - disassemble after the nucleoside cofactor - in this case ATPγS - is hydrolysed (see figure 6.6). Indeed it has been shown by others, to take between 1.5 - 2 hours for all NPFs produced with ATPγS to disassemble from a population.[183]

Therefore, NPFs produced with ATPγS are considered stable able to form multiple synaptic joints – which are suggested to last approximately 5 minutes for a 60 nt NPF[17] – concurrently.

However, the population of synaptic joints is actually observed to decrease throughout 60 minutes in these experiments. Given the discussed disassembly rate, it is possible that a large proportion of the NPF population begins to disassemble between 30 - 60 minutes.

This would explain a overall reduction to approximately 50% of the DNA population indicating an ongoing nucleoprotein interaction. It is important to note that the proportions of synaptic and postsynaptic complexes within this population is seen to vary as a consequence of NPF MW.

However, such loses due to the turnover of ATPγS would be expected to be a function of temperature, yet similar NPF populations are observed at $22°C$ and $37°C$, which cannot be explained here.

Moreover, the occurrence of parallel synaptic joint formations is reconciled by the observation of NPF clusters. These are resolved throughout the studies in this thesis suggesting that the cooperation between NPFs acts to increase the local concentration in the vicinity of the DNA template, thus facilitating a more efficient interaction through the confinement of reactive species.

As such, it may be proposed that this cooperative search reduces the time required to probe the available sequence space in this study by a typical factor of 3 - where an average of three synaptic joints were observed on a 890 bp DNA at any one time.

In reality any efficiency saving will directly relate to the DNA template length, complexity and crucially; degrees of freedom. Nonetheless, it is seen that up to half of the available sequence space can be simultaneously occupied and in some cases directly juxtaposed joints can exist on the 890 bp templates used here. This would suggest that in theory the entire template could be occupied, as there appears to be little resistance to synaptic complexes existing shoulder-to-shoulder.

Interestingly, it is possible that the accumulation of NPF clusters may sequester the interacting NPF population over time resulting in a net reduction in the observed DNA population undergoing a nucleoprotein interaction. This may account for the discrepancy in temperature dependence of the diminishing NPF population, where cluster formation would be expected to be favoured with reduced temperature, in contrast to ATPγS turnover. Further work would be required to experimentally verify if the occurrence of NPF clusters increases consistently over time.

This page is intentionally left blank.

Chapter 7

Tuning the Translational Freedom of DNA for High Speed AFM

7.1 Introduction

To date, interrogation and application of the RecA mediating homology search has been conducted primarily through solution based biochemistries. Such approaches provide an environment that closely mimics native physiology, yet limit the resolution of applied interrogation methodologies. Such an approach was demonstrated in the previous chapter, where the interaction was conducted in solution, but analysed dried on a surface with the AFM.

In contrast, the work in this thesis aims to apply the high spatial and temporal resolutions of HS-AFM to directly interrogate the interaction of NPFs with dsDNA. As, HS-AFM utilises a physical probe, such interactions must take place at a solid–liquid interface in order to be observed (see section 4.2).

However, the confinement of interaction species in a two dimensional plane may irrecoverably perturb the interaction of interest, where the degrees of freedom of the subject dsDNA molecule have been shown to be of critical importance.[11] This was further evidenced by work in the previous chapter (section 6.2.5) by the break down of the resolution phase of the reaction during surface experiments.

With respect to this, it is important to mediate a transient absorption of dsDNA substrates upon a surface such that a pseudo three dimensional interaction may be orchestrated. The work in this chapter investigates sample preparations to tune the surface association – and therefore degrees

of freedom – of dsDNA, applicable to conducting NPF homology searching interactions at a solid–liquid interface. Moreover, such control over the surface association and molecule freedom are likely to be of critical importance in interfacing DNA–organised devices with macroscale components in the future.

7.1.1 Mediating the Surface Interaction of DNA

The sample preparation procedure and its impact on the experiment is a key concern in all forms of microscopy. Here, AFM has distinct advantages over radiative based alternatives as it does not require the labelling or modification of delicate biological structures. It does, however, employ a physical probe (section 4.1) and therefore requires the deposition of the sample on a solid support surface.[210] Muscovite mica surfaces are typically used as a substrate owing to their perfect cleavage along a $< 001 >$ plane, yielding large atomically flat areas. Mica consists of layers of an aluminium phyllosilicate lattice ionically bonded through interstitial K^+ions. It is well established that upon cleavage these K^+ions are highly mobile and are readily exchanged with divalent cation species at the solid-liquid interface.[324] This exchange results in a positive overcharging of the mica surface, which enables the deposition of molecules that hold a net negative charge, such as DNA.[325]

However, the 2D confinement of the molecules inherently perturbs any biological interaction of interest, as highlighted in the previous chapter (section 6.2.5). It is therefore necessary to mediate a surface absorption which allows the retention of sufficient translational and rotational freedom whilst securely adhering the molecules against the instantaneous lateral forces imposed by the scanning probe. Additionally, it is necessary to maintain a suitable reduced physiological environment for biological activity to be retained.

Previous studies employing methods to modify the mica surface chemically,[326] and the use of cyclic dielectric fields[327,328] and cyclic buffer exchanges,[329,330,331] have attempted to mitigate confinement effects by cycling the surface between a loose association - allowing reactions to proceed freely - and a tight association to enable high-resolution imaging. It has also been shown that modulation of surface voltage can be used to control adhesion forces,[332] but this requires the use of metal electrodes as substrates which typically confer poor surface roughness. These alternative methods either put significant constraints on the experimental setup or do not provide the required functionality, and therefore pre-incubation of mica with divalent cations remains the common practice for biological samples.

It is well established that upon cleavage of mica layers the interstitial K^+ions are split between the two cleavage faces and tend to cluster on the resultant surface.[333] The resulting vacant sites are easily filled upon exposure to divalent cations. Transition metal cations, such as Ni^{2+}, $Co^{2+}$and Zn^{2+}bind irreversibly to mica, and have also been demonstrated to displace residual K^+ions. They interact with the recessed hydroxyl group of the mica[325,333] cooperatively, leading to phase separated domains of positive surface charge.[319] Where transition metal cations are utilised, a strong adhesion between DNA and mica is achieved due to directional bonding of the d-orbitals of the divalent cation. Such deposited DNA molecules adopt a kinetically trapped form,[319] sufficient for high-resolution imaging in aqueous buffer.

In contrast, Mg^{2+}cannot form directional bonds, continually exchanging with residual $K^+$and H^+ions on the mica surface and in the imaging buffer. As a consequence, Mg^{2+}only mediates a relatively weak and diffuse adhesion between DNA and the mica surface, allowing DNA molecules to adopt an equilibrated form.[325,319,334,335] This complex spatial distribution has been rationalised by others,[319,335] where the phase separated domains of ionic species are represented as an idealised polka dot pattern - an ionic patchwork of evenly distributed positively charged Ni^{2+}clusters interspersed across a mica surface undergoing dynamic exchanges between K^+, $H^+$and potentially other cations such as Mg^{2+}.

Previous work by Pastre et al, has taken advantage of this difference in binding capacities by adjusting the ratio between different monovalent ion and monovalent to divalent ion concentrations in the deposition buffer to adjust the surface adhesion of DNA in air.[230,336,337] Interestingly, $Mg^{2+}$is the primary co-factor for the majority of enzymes that manipulate all biological polyphosphates including ATP, DNA and RNA.[338] Thus, the transient nature of Mg^{2+}association allows it to be utilised, in equilibrium, as both surface binding agent and enzymatic co-factor. Typically, cellular Mg^{2+}ranges from 5 – 20 mM, where 95% is chelated with proteins and nucleic acids, primarily ATP.[338] However, a concentration below 10 mM is sufficient for most nucleoprotein interactions. In contrast, Ni^{2+}exhibits little to no physiological role in higher organisms, enabling its use as a strong surface binding agent.

Building on the work by Pastre et al[337], this chapter explores the use of monovalent and divalent cations to mediate the continual surface attachment of dsDNA templates throughout HS-AFM imaging. Initially, a re-examination of the critical transition region between strong and weak absorption - demonstrated in the original study[337] - is undertaken. This is subsequently extended to imaging DNA dynamics in aqueous buffer. Further, development examines the modulation of divalent – divalent cation ratios to tune the surface translational freedom of DNA molecules,

implemented alongside a novel methodology for quantifying the DNA mobility.[320] It is important to acknowledge the contribution of Dr Mike Szymonik to the coding of this algorithm. Finally, the method is demonstrated to enable real time observation of protein – DNA interactions with the HS-AFM. It is important to note that this text is derived from work published by the author. For the original text please see Lee et al[320]

7.2 Experimental

7.2.1 Re-visiting Pastre's Transition Zone

Initial work to mediate a transient interaction between dsDNA and the mica substrate, began by re-visiting the work conducted by Pastre et al.[337] Here, by manipulating the ratio of Na^+ and Mg^{2+}in the deposition buffer, Pastre et al demonstrated it was possible to increase or decrease the deposition of dsDNA on mica.

From this work a transition zone between surface adsorption and non-adsorption (figure 7.1 A, grey area) was identified, where the prevalence of one state was defined by the presence of over 50% of the DNA molecules.[337]

As the aim of this chapter is to mediate a transient adsorption of DNA, this transition zone was reasoned as the perfect place to start. A total of 15 ng of 890 bp dsDNA was deposited on freshly cleaved mica in a range of 10 mM Tris-HCl (pH 7.5) buffers (methods 5.1.3). Six buffers were examined, containing ratios of mono and divalent cations ranging from 5 to 15 mM $Mg^{2+}$and from 50 to 150 mM Na^+ - these are identified in figure 7.1 A (orange arrows). DNA molecules were deposited for 5 minutes prior to rinsing with a volume of the same buffer and drying (methods 5.2.8). Substrates were imaged at random locations and the representative distribution of DNA molecules was counted.

Figure 7.1 shows representative images from the above experiments. From these, it is clear that where the concentration of $Mg^{2+}$is modulated from 15 - 5 mM (figure 7.1 C - E) - relative to a constant 100 mM Na^+ - the number of surface bound DNA molecules drops off rapidly, from 84 $\pm$ 3 to only 2 $\pm$ 2 per 6.25 μm^2, respectively. This is interesting, as this represents a decrease in the estimated surface coverage from 1 ng to 24 pg of the original 15 ng of dsDNA deposited, far from the 50% suggested by Pastre et al.

In contrast, moving from 50 to 150 mM Na^+ (figure 7.1 F - H) - relative to a constant 10 mM

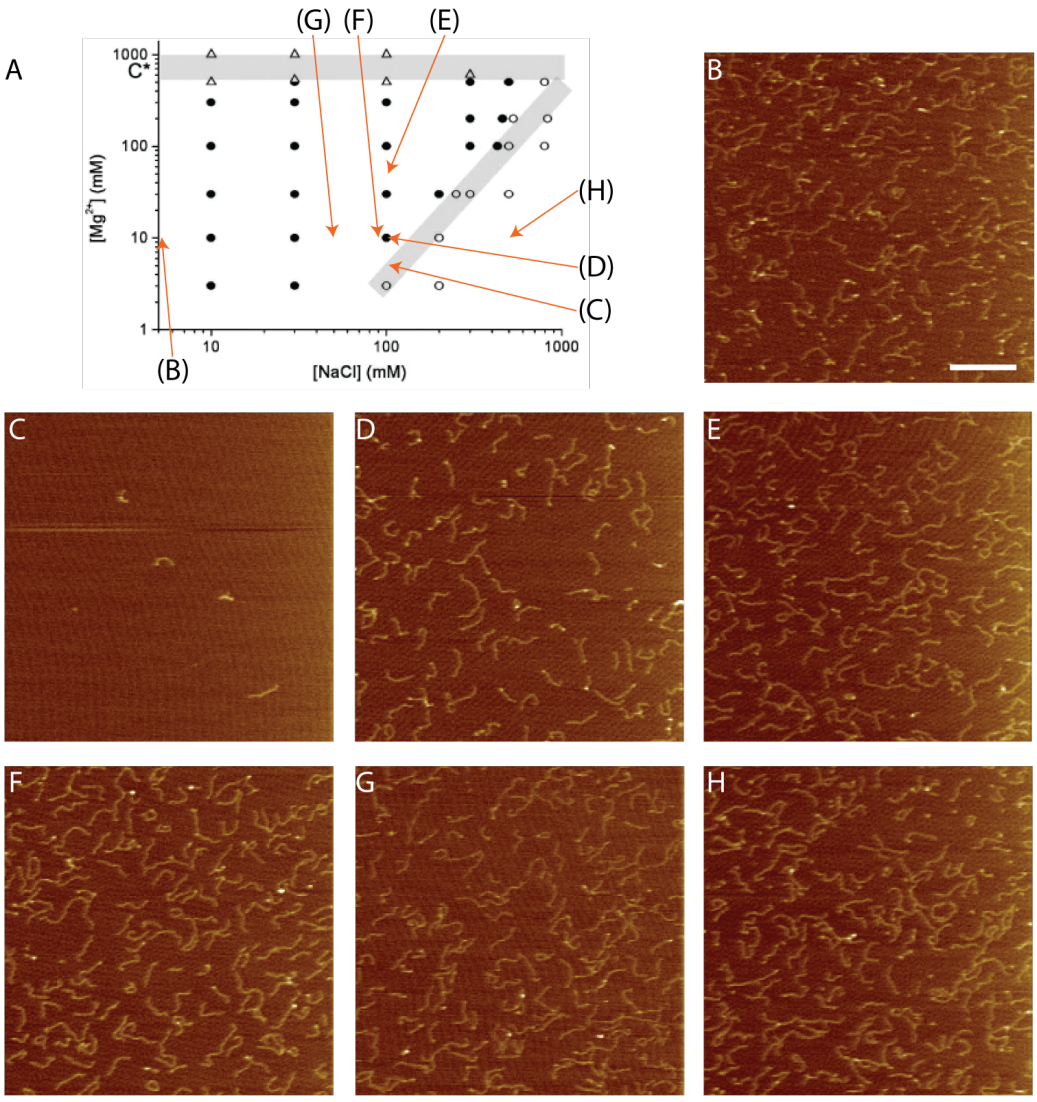

Figure 7.1: ***The re-investigation of a binding transition zone as originally investigated by Pastre et al[337].*** *AFM images taken in air of the resultant DNA surface population when deposited in buffer containing the ratio of Mg^{2+}:Na^{+} indicated by the plot in (A). B) 10 mM Mg^{2+}control. C) 5 mM:100 mM, D) 10 mM:100 mM, E) 15 mM:100 mM, F) 10 mM:100 mM, G) 10 mM:50 mM and H) 10 mM:150 mM. Z scale = 3 nm, Scale bar = 500 nm. How the cation ratios studied here correspond to the transition zone identified in Pastre's work (shown in grey) is indicated by the orange arrows on plot (A). The plot in (A) is adapted from Pastre et al.[337]*

.

Mg^{2+}- was found to have little effect on dsDNA deposition, contrary to that suggested by Pastre (figure 7.1 A), with a typical surface coverage of 88 $\pm$ 4 DNA molecules per 6.25 μm^2. This does, however, represent a departure from the control deposited in 10 mM Mg^{2+}only, where 113 $\pm$ 9 DNA molecules per 6.25 μm^2 were observed. This difference is reasoned to be due to the formation of a electrostatic double layer at the mica surface by the Na^{+}, causing a neutralisation of the surface charge and a consequential reduction in Mg^{2+}association.

From these experiments it becomes evident that the transition zone is slightly shifted from that previously identified. This is critical as, due to the linear nature of the nucleic acid polyelectrolyte, its charge density scales with length. Hence, one would expect the transition zone to scale with respect to the length of the DNA molecule utilised.

Here the transition is noted from 10 mM Mg^{2+}decreasing through to almost complete non-adsorption at 5 mM Mg^{2+}(figure 7.1). When it is considered that 10 mM $Mg^{2+}$is also a known requirement of the RecA:dsDNA interaction then it was argued that this stipulates an optimal Mg^{2+}concentration for any future experiments.

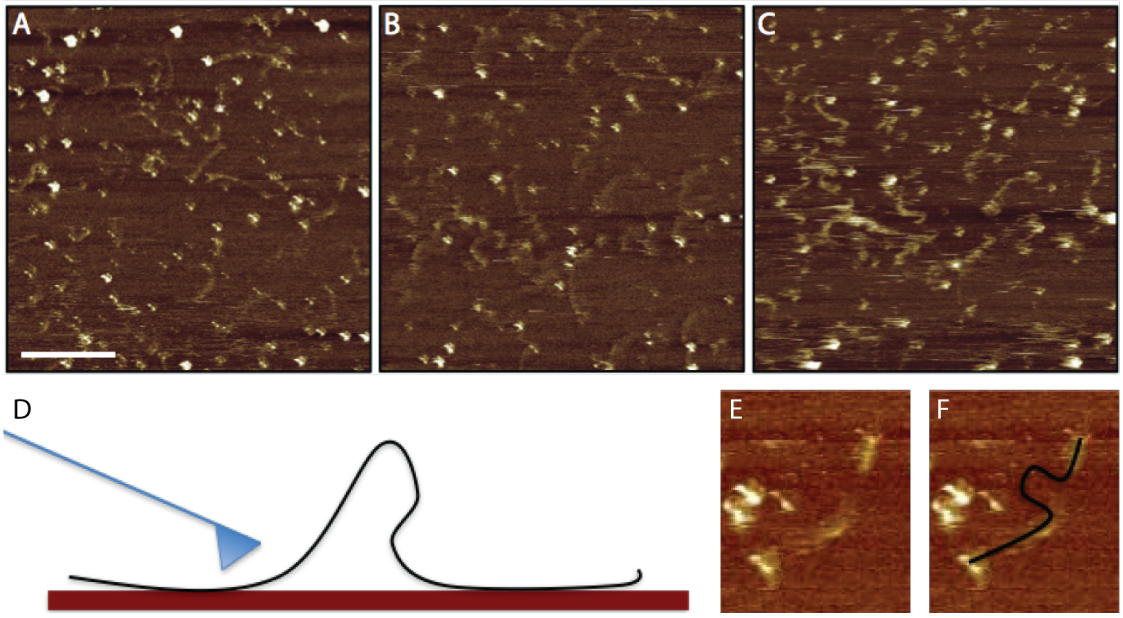

Figure 7.2: ***The effect of screening charge on 890 bp observations.*** ***(A - C)*** *Representative AFM images taken in aqueous buffer consisting of 10mM Mg^{2+} and 100 mM Na^{+}. 890 bp DNA is faintly resolved due to net repulsion of the tip from the extensive electrostatic double layer formed by condensation of Na+ to the mica surface is seen to obscure the spatial resolution.* ***(D)*** *A schematic diagram depicting the partial binding of 890 bp DNA leading to the resolution of incomplete molecules as indicated in* ***(E)****. The suggested molecule conformation is superimposed in* ***F****. Z scale = 8 nm, scale bar = 250 nm.*

As such, a buffer consisting of 10 mM $Mg^{2+}$and 100 mM Na^{+} was utilised in an attempt to

orchestrate a transient surface binding of DNA during dynamic observations. 890 bp DNA was deposited as previous without drying and imaged continuously within the same buffer. Representative images are depicted in figure 7.2, from which it can clearly be seen that 890 bp DNA molecules are poorly resolved. This is a result of the significant tip repulsion caused by the large amounts of Na^{+} that concentrates at the surface of the mica and around the AFM probe. This causes an electrostatic double layer (EDL) (see section 4.2.1) that extends far away from the mica surface, repelling the probe as it pushes through at the extend of each cantilever osculation, representing a much larger energy dissipation.

Nonetheless, the DNA molecules appear to be highly mobile under these conditions. Interestingly, most of these molecules were measured to be considerably shorter than the 302 nm expected. This can clearly be seen in figure 7.2 A - C, with reference to the scale bar. Although it is possible that the DNA has been truncated in some way, it is more likely that the appearance of short species is evidence of partially bound regions of the constituent molecules as depicted in figure 7.2 D - F. When taken together, these observations suggest that the DNA molecules are very loosely associated with the surface, largely in equilibrium between being solvated in the buffer and transiently present at the surface.

7.2.2 Divalent–Divalent Cation Ratios

The severe reduction in spatial resolution was considered an insurmountable compromise and hence the utilisation of monovalent cations concluded to be inappropriate for this work. In contrast, previous work by the author and others has explored the binding capacity of Ni^{2+}, demonstrating this transition metal cation's ability to mediate a strong surface adsorption.[319]

Here, employing the previously discussed, idealised polka dot representation (section 7.1.1), it was argued that by pre-incubating the mica surface with different concentrations of Ni^{2+}, it would be possible to tune the extent of strong localised surface-DNA association in line with the Ni^{2+}concentration, where DNA molecules are deposited in Mg^{2+}containing buffer. Here, $Ni^{2+}$was reasoned to restrain the translational freedom through a few specific anchor points whilst Mg^{2+}should mediate a more global and distributed, but weak association.[325,319,334,335] Thus the mobility of a DNA molecule maybe considered a function of the number of Ni^{2+}interactions - controllable by the Ni^{2+}concentration and exposure time during surface pre-incubation.

Initial experiments indicated this to be a valid approach, with highly mobile 890 bp dsDNA observed across 5 - 15 mM Ni^{2+}incubated mica substrates, when deposited in a 10 mM Tris-

Mg^{2+}buffer (methods 5.1.3). However, the DNA molecules were found to be too mobile, observed to undergo large "worm-like" movements across the entire imaging area, even at the higher Ni^{2+}concentrations. This can clearly be seen in figure 7.3 & movie 7.1 for 10 mM Ni^{2+}.

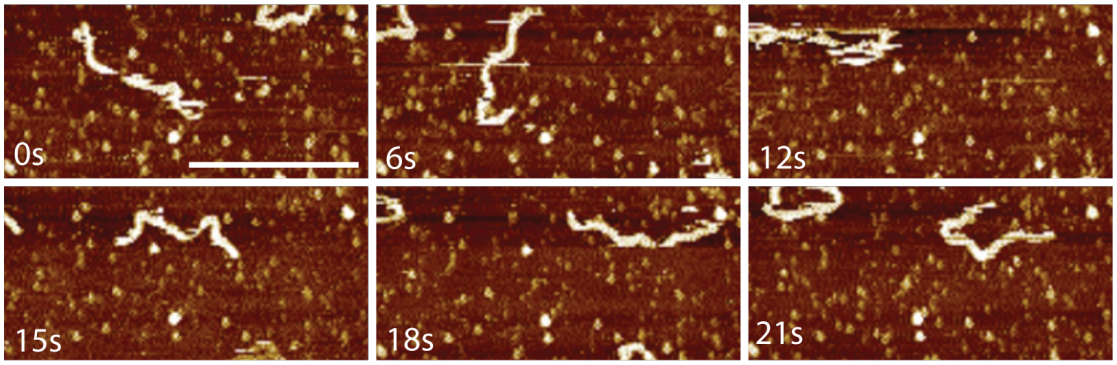

Figure 7.3: ***Surface translational mobility of 890 bp DNA molecules.*** *890 bp DNA molecules are observed to undergo global translational movements around the imaging window when immobilised on a 10 mM Ni^{2+}pre-incubated mica substrate. Scale bar = 250 nm. Z scale = 3 nm.*

This is an important point to consider, as it is necessary to continuously image individual molecules over the time course of the RecA:DNA reaction. If the surface association is too weak, the DNA molecule may be able to undergo global translational movement out of the imaging frame causing key events to be missed. Optimal binding conditions therefore, need to present weak surface association without large global translational movements of the DNA across the mica surface, while at the same time allowing local translation. As already stated, the charge density of DNA is a function of length, therefore, the mobility decreases in-line with the increasing molecule length under any given surface preparation. With respect to this, 3.5 kbp dsDNA was found to confer local but limited global translational movements from image to image when observed under the same conditions as the previous 890 bp. It was concluded that utilisation of a larger DNA molecule was a preferably alternative to drastically increasing the Ni^{2+}– beyond 15 mM – which typically resulted in detrimental surface crystallisation. Moreover, in light of the DNA molecules observed degrees of freedom it was concluded that a methodology for quantifying the surface translational movements was necessary in order to identify preferential surface associations.

7.2.3 Quantifying the Mobility of DNA Molecules

In order to establish a methodology for quantifying the mobility of surface-immobilised DNA molecules, mica substrates were pre-incubated with 15 mM Ni^{2+}. Subsequently, 3.5 kbp DNA

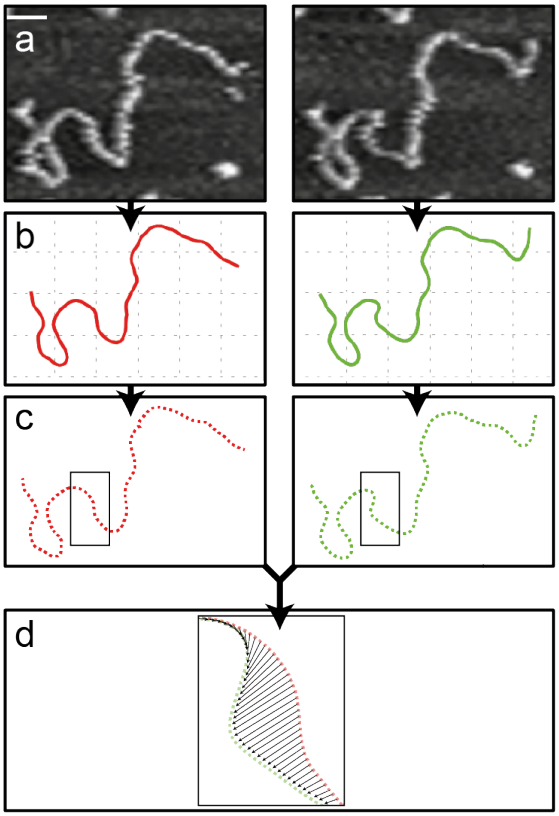

Figure 7.4: ***A schematic diagram depicting the chain fitting and evolution processing from the raw AFM images.*** *(**a**) Sequential AFM images following a single DNA molecule. (**b**) Derived chain vectors. (**c**) Mapping of chain vectors into equidistant points. (**d**) Point mapping is indicated across a short section of the molecule; the relevant section of the molecule is indicated by the boxes in (**c**). The scale bar is 100 nm.*

molecules were deposited in Tris-Mg^{2+}buffer (methods 5.1.3) and then imaged by AFM. 15 mM $Ni^{2+}$was expected to provide a strong surface adhesion, allowing contrasting small regions of chain mobility to be easily identified. 1 μm^2 areas of the sample were continuously imaged in solution to observe the motion of the DNA molecules. For each sequence of AFM images, several representative DNA molecules were selected and processed using a chain fitting algorithm, derived from the open source Jfilament ImageJ plug-in.[339] A chain of co-ordinates was fitted to the height peaks in the AFM image, generating a vector representation of the molecule's conformation (figure 7.4 (a) & (b)). This method eliminates the noise and image artefacts that arise during HS-AFM imaging in liquid. Each selected DNA molecule was tracked sequentially through an image series, where fitting was repeated for sections of chain that deviated from the previous image, allowing their motion to be quantified. Each chain was then divided into 100 equidistant points and a minimisation algorithm was used to map corresponding points between successive frames in a time series, mitigating errors in imaging and fitting around the chain ends. The amount of motion for each pair of points could then be calculated, averaged and normalised to yield movement velocities for individual DNA chains or sections of a chain (figure 7.4 (c) & (d)). An overview of this methodology is given in movie 7.2.

To quantify the lateral motion of each DNA molecule over time, the movements of individual

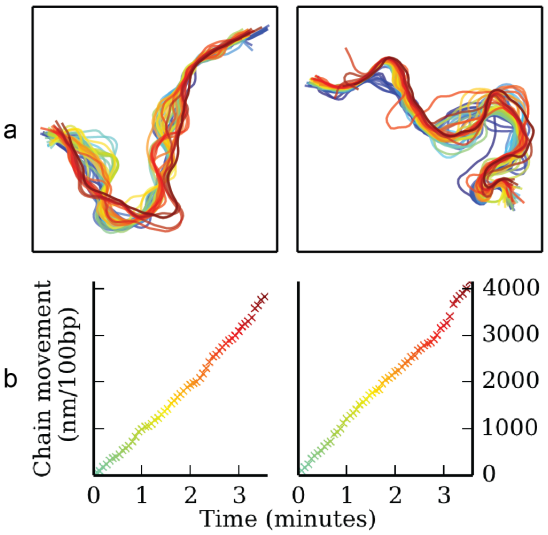

Figure 7.5: ***Sequential chain mapping and movement of two representative DNA molecules.*** ***(a)*** *Sequential HS-AFM images of two representative 3.5 kbp DNA molecules over long time scales at surface pre-incubation Ni^{2+}concentrations of 15 mM. The time evolution of the DNA chain is represented by a colour scale, with the lightest blue line representing the initial state, and the darkest red line the final state. Under these binding conditions DNA molecules have little mobility indicated by the tight overlap through subsequent frames. Panel* ***(b)*** *shows the chain movement per 100 bp for both DNA molecules, from which an average chain mobility of 17 $nm100bp^{-1}s^{-1}$ can be calculated.*

molecules were analysed. Figure 7.5 (a) shows the time-course of two representative DNA molecules over the course of an image sequence. The time evolution is represented by a colour scale, with the darkest red line representing the final state, and the lightest blue line the initial state. The total distance travelled by the DNA chain, normalised by its length, was determined as the sum of all the distances travelled by the equidistant points on the chain. The results for the two representative DNA molecules are shown as a function of time in figure 7.5 (b). The movements were found to be linear in time and hence, from the plots in panel (b), a velocity per unit of DNA in nanometres per 100 base pairs length per second (nm 100 bp^{-1} s^{-1}) was then derived as the characteristic quantity to measure lateral mobility.

Analysing the motion of DNA molecules is not a novel concept. Previous work has attempted the direct superposition of sequential images aligned to immobile parts of the molecule[340] or to a reference elsewhere in the image.[341] Deriving additional detail by analysis of vectorised molecules has also been undertaken through hand-tracing methods[342] to examine quantities such as elastic bending energies[343] and the influence of the AFM probe.[344] However, such studies lack detail due to the crude alignment and vector derivation methodologies, yielding mobility analysis based upon single point mapping of centre-of-mass or chain termini only. Consequently, these analyses lack the power to provide the understanding necessary to tune the translational freedom of DNA to enable relevant nucleoprotein interactions to take place.

7.2.4 Tuning DNA Surface Mobility With Ni^{2+} Pre-incubation

To investigate whether the DNA mobility can be tuned as a function of surface-immobilised Ni^{2+} ion density, the above approach was applied across a range of Ni^{2+} concentrations from 1 mM to 15 mM (figure 7.6). Varying the ion concentration was chosen over incubation time, which may lead to unwanted crystallisation of Ni-salts for longer exposure times. A lower Ni^{2+} concentration is expected to lead to a lower density of positively charged sites on the surface, providing fewer pinning points for a DNA molecule and allowing it more translational freedom.

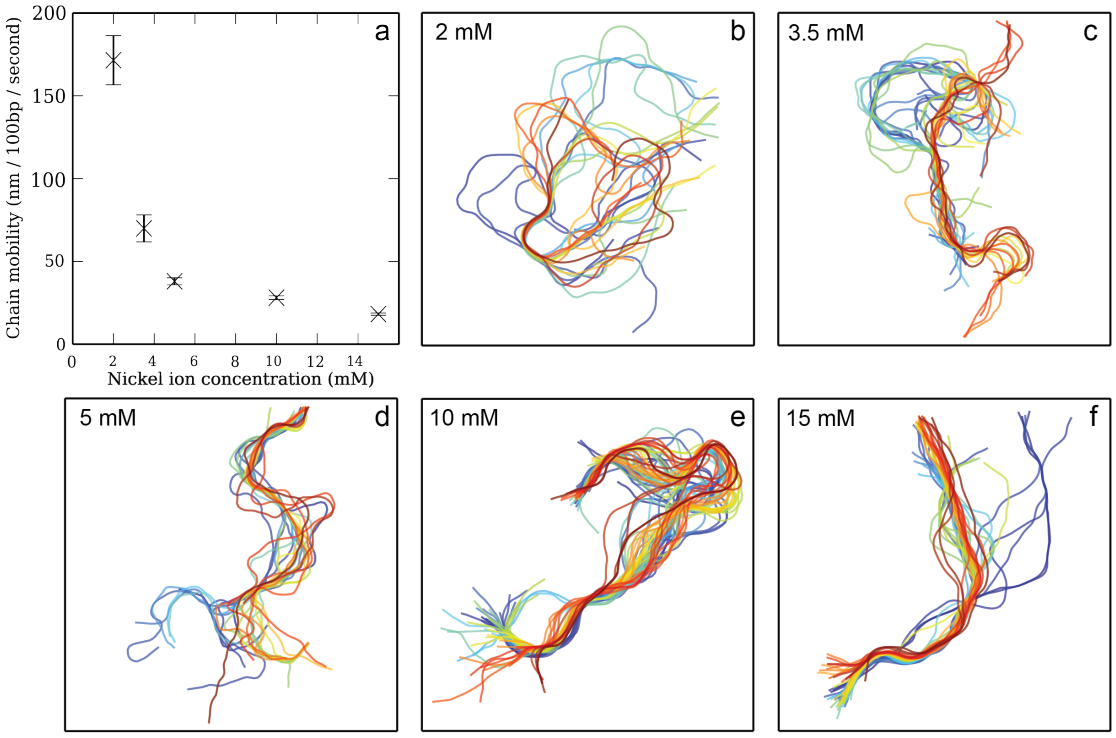

Figure 7.6: ***DNA molecule translational mobility*** *($nm100bp^{-1}s^{-1}$)* ***as a function of Ni^{2+} surface pre-incubation concentration.*** ***(a)*** *DNA molecule translational mobility ($nm100bp^{-1}s^{-1}$) as a function of Ni^{2+} surface pre-incubation concentration.* ***(a)*** *A distinct decrease in mobility is associated with an increasing Ni^{2+} concentration. Error bars indicate the standard deviation for each data point. Example chain vector evolution profiles for* ***(b)*** *2 mM,* ***(c)*** *3.5 mM,* ***(d)*** *5 mM,* ***(e)*** *10 mM, and* ***(f)*** *15 mM Ni^{2+} surface pre-incubation concentrations.*

Figure 7.6 (a) shows the chain mobility as a function of Ni^{2+} concentration (also see movie 7.3). Each point on the plot represents an average taken across a number of DNA molecules and the error bars represent the standard deviation. It can be seen that the DNA translational mobility increases significantly with decreasing Ni^{2+} concentration. Where the surface concentration of Ni^{2+} is at the highest level investigated here, large regions of the chain remain stationary through

many consecutive images in a series, indicating very tight surface association (see figure 7.6 (f) for an example).

In contrast, below 5 mM the surface Ni^{2+}becomes sufficiently sparse that large stretches of the DNA chain show considerable movement across the time frame of the experiment (figure 7.6 (d)). This is also reflected in the chain mobility, which was calculated to be 38 $\pm$ 4.5 nm 100 bp^{-1} s^{-1} at 5 mM Ni^{2+}, with even larger degrees of freedom found for 3.5 mM and 2 mM Ni^{2+}(70 $\pm$ 8 nm 100 bp^{-1} s^{-1} and 171 $\pm$ 15 nm 100 bp^{-1} s^{-1}, respectively). For the lowest Ni^{2+}concentration investigated (1 mM), a high velocity of motion was observed (see movie 7.3) which meant that a meaningful quantification of the mobility was not possible, so this concentration is not represented in the plot.

From figure 7.6 it can be concluded that for the 3.5 kbp DNA molecules used in this study, optimal conditions are achieved at around 5 mM - 10 mM Ni^{2+}, producing optimal spatial resolution and chain flexibility, with a mobility of 38 $\pm$ 4.5 - 28 $\pm$ 2.4 nm 100 bp^{-1} s^{-1}. This is proposed to be sufficient for imaging enzymatic interaction with minimal hindrance from the mica support surface, whilst the DNA molecules remain relatively immobile within a 1 μm^2 imaging area.

It is important to note that, as discussed previously (section 7.2.2), the values quoted here are normalised with respect to the length of the DNA molecule. Therefore, while a surface preparation can be tuned to produce the desired degree of nucleic acid association, it is necessary to adjust conditions to account for the length and charge of the particular molecules of study.

7.2.5 Influence of the HS-AFM Probe

When imaging biological systems, the tip radius of the AFM probe is generally of the same order of magnitude as the surface-immobilised molecules, and hence it can be expected that the interaction between the two is significant. Furthermore, in the investigation reported here, the DNA molecules are deliberately only weakly associated with the surface, leading to increased impact from the tip – molecule interactions. The vectorisation of chain motion employed here enables the behavioural analysis of individual segments of each DNA molecule, and hence allows for the tip – DNA interaction to be analysed.

As the direction of movement for each DNA chain segment is known, a polar histogram of the magnitude of each movement versus the direction can be plotted for the whole dataset across all Ni^{2+}concentrations (figure 7.7 (a)). A distinct preference for movements along a particular direction can be observed, where an anisotropy of 1.4 was obtained by fitting an ellipsoid to the

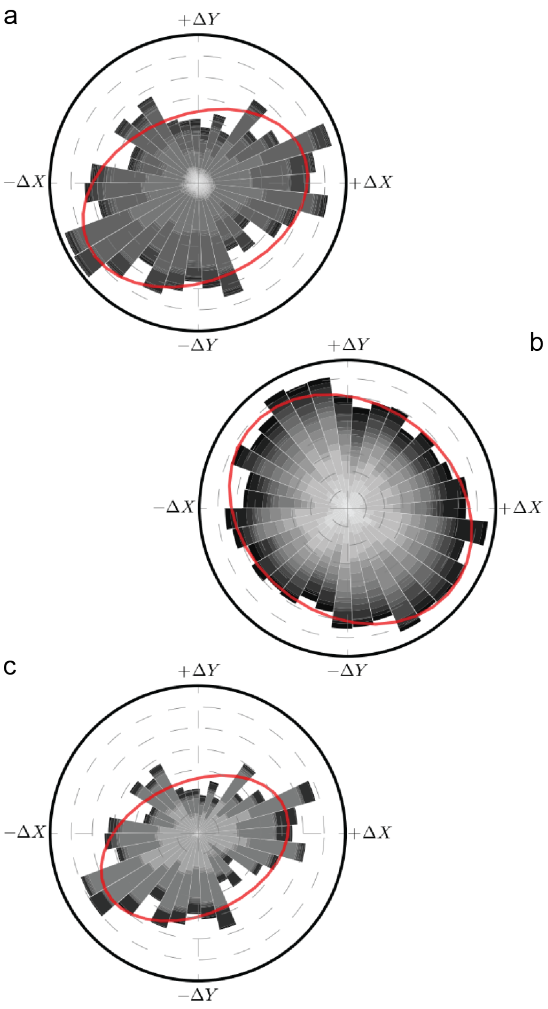

Figure 7.7: ***Polar histograms indicating the magnitude and directionality of chain movements.*** ***(a)*** *including all Ni^{2+}concentrations,* ***(b)*** *including only high Ni^{2+}concentrations (10 mM and 15 mM), and* ***(c)*** *including only low Ni^{2+}concentration (2 mM and 3.5 mM). The movements of all segments are binned according to the direction and weighted by the magnitude of the movement. The grey scale delineates the contributions of different DNA chains. The degree of the anisotropy of each histogram is indicated by the elliptical fit to the data (indicated by the red ellipsoid).*

data. Random thermal motion of the DNA molecule is not expected to lead to any directional preference. In contrast, the AFM tip is scanned back and forth very fast along one axis (the fast scan axis), and relatively slowly along the other axis. Therefore, the tip - molecule interactions are expected to predominate in the direction of the fast scan axis, and hence would lead to an anisotropic histogram.

When examined more closely, it can be seen that the anisotropy arises predominately from highly mobile DNA molecules that appear on low Ni^{2+}sample preparations. Figure 7.7 (b) shows a polar histogram of the magnitude of the movement versus the direction but with only the Ni^{2+}concentrations of 10 mM and 15 mM included. From the fit of the ellipsoid to the data, an anisotropy of only 1.2 is obtained. In contrast, when considering only the very weakly bound DNA molecules, which show significant lateral mobility (2 mM and 3.5 mM Ni^{2+}, figure 7.7 (c)), the anisotropy is found to be more pronounced than in panel (a) (anisotropy of 1.5), demonstrating that the anisotropy increases with increasing lateral mobility.

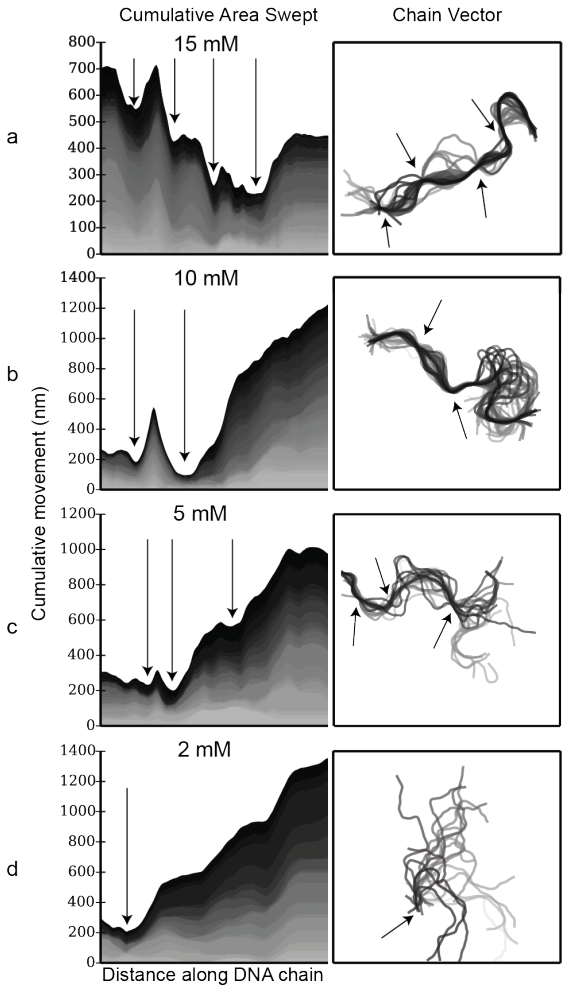

Figure 7.8: ***Area plots depicting the cumulative movement of chain segments for representative DNA chains across different Ni^{2+} concentrations.*** *Distinct regions of the DNA molecule are seen to be anchored at particular locations (indicated by arrows). The density of minima increases with Ni^{2+} concentrations, suggesting that Ni^{2+} ions provide specific anchor points on the surface to increase DNA surface association. The grey scale indicates successive frames in a time series. The panels on the right show the chain evolution with the minima location indicated by arrows.*

Notably, the skew seen in the magnitude data is found to be in line with the image acquisition direction, suggesting that the AFM probe contributes significantly but not overwhelmingly to the observed motion of loosely bound DNA molecules. Conversely, higher Ni^{2+} preparations are characterised by smaller magnitude motions and are largely isotropic in nature. This would be consistent with molecules undergoing small conformational changes as a result of thermal fluctuations, notably constrained against large translational probe-induced motions.

As observed, the lateral forces imparted by the AFM probe are likely to perturb the sample of interest and may even influence any nucleoprotein interactions. However, the ability to extract detailed information regarding the behaviour of these molecules during the imaging process may enable the deconvolution of these contributions from observed biological interactions in the future.

7.2.6 Identification of Discrete Surface Anchor Points

To consider the role of Ni^{2+}within this system, the idealised polka dot surface model discussed by others was adopted.[319,333,334] In this work it was reasoned that the density of surface-associated Ni^{2+}ions increases with increasing Ni^{2+}concentration during incubation. It has previously been demonstrated that the distribution of Ni^{2+}on the surface is highly non-uniform[343] and lacks homogeneity.[333] Indeed this is represented in the increased mobility variation observed for lower Ni^{2+}concentrations. However, it may be possible to gain further insight into the Ni^{2+}/Mg^{2+}governed surface interactions by identifying regions of specific surface association through examination of the motion of DNA molecules.

Plotting the cumulative movement for each segment along individual DNA molecules (figure 7.8), a series of minima can be observed (indicated by arrows). Figure 7.8 (a) shows the results for a representative DNA molecule immobilised on a 15 mM Ni^{2+}pre-incubated surface, and four distinct minima can be identified. The right hand panel shows the time evolution of the DNA molecule, from which particular regions of the chain appear pinned to the surface while other regions move large distances. The pinned, stationary positions coincide with the observed minima in movement, and are again indicated by arrows.

The number of such minima, and hence the number of stationary points, decreases with decreasing Ni^{2+}concentration (figure 7.8 (b)-(d)). It should be noted that all of the molecules imaged on surfaces pre-incubated with 10 mM $Ni^{2+}$and in excess of half of the molecules imaged on 5 mM Ni^{2+}surfaces show distinct pinning at one or more locations. Interestingly, in some cases minima disappear over the course of the imaging sequence, with others appearing in subsequent frames, suggesting that specific surface interactions break and re-form continuously; this behaviour is more prominent with molecules imaged at 5 mM or lower Ni^{2+}concentration. Although a direct correlation with surface distributions of Ni^{2+}would be purely speculative without further in depth studies, the results together with previous findings on the distribution of the Ni^{2+}ions[333,343,319] suggest that positively charged Ni^{2+}domains act as anchor points for the negatively charged DNA chains - with the interstitial regions governed by weak Mg^{2+}associations - allowing the desired mobility.[320]

7.2.7 In Situ Digestion of DNA by EcoRI

In order to demonstrate the utility of this approach for the observation of nucleoprotein interactions, EcoRI was chosen as a proof-of-principle for *in situ* enzymatic action. A restriction

enzyme offers an elegant demonstration of protein function, yielding the clearly discernible and unambiguous result of generating a short DNA fragment cut from the original 3.5 kbp DNA molecule.

3.5 kbp DNA was immobilised on 5 mM and 15 mM Ni^{2+} pre-incubated mica surfaces and imaged in Tris-Mg buffer (methods 5.1.3) as discussed in section (7.2.4). Stable imaging was established prior to the introduction of 0.5 units of EcoRI. The reaction was allowed to proceed at room temperature for 1 hour whilst continually imaging throughout.

Final statistics taken after 1 hour of incubation indicate that EcoRI is able to digest up to 43% of the DNA molecules present on a 5 mM Ni^{2+} surface, compared to only 6% on a equivalent 15 mM Ni^{2+} substrate (figure 7.9). When taken together with observations of the respective surface translational mobility (figure 7.6), this suggests that tightly bound molecules inhibit the ability to observe DNA protein interactions *in situ* with the HS-AFM.

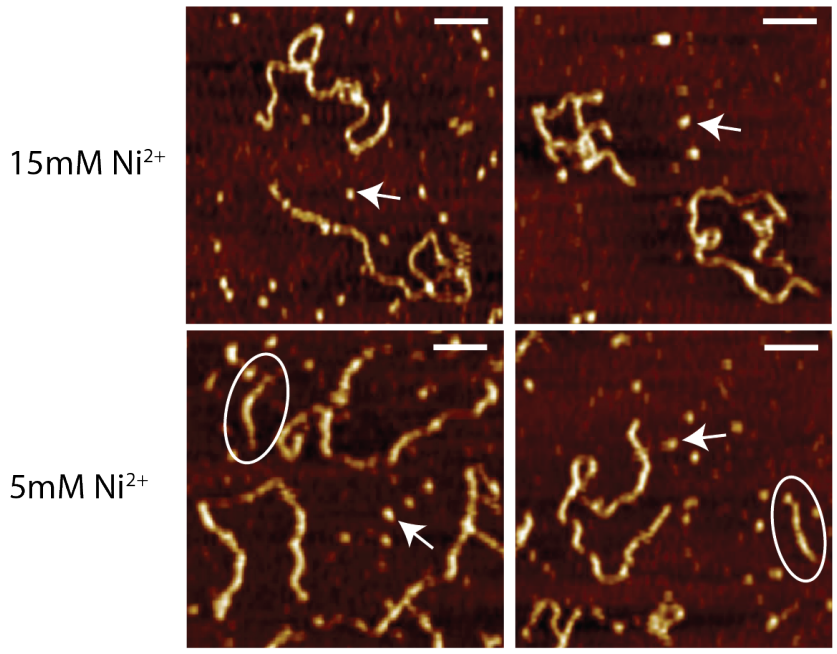

Figure 7.9: ***HS-AFM observations of EcoRI restriction digestion of 3.5 kbp DNA immobilised on 5 and 15 mM Ni^{2+} mica surfaces.*** *On the 15 mM Ni^{2+} treated mica surface the DNA molecules are seen to be tightly bound, adopting a kinetically trapped configuration. In contrast the DNA molecules are observed to have significant mobility on the 5 mM Ni^{2+} treated mica surface, resulting in successful restriction digestion. Digestion fragments are circled in white, and EcoRI proteins are indicated by white arrows. Scale bars are 125 nm. Z scales are 5 nm.*

Moreover, the digestion of DNA molecules can be observed directly at the single molecule level, clearly discerning the excision of a short DNA fragment on a 5 mM Ni^{2+} surface (figure 7.10

and movie 7.4). Due to the relatively small size of the EcoRI enzyme (31 kDa) it is difficult to discern individual proteins arriving alongside the mobile DNA molecule. However, the action of the EcoRI restriction enzyme is clearly evident, as the DNA is separated into two molecules (figure 7.10 12s). The observation of these newly generated separate molecules (indicated by a star) unambiguously indicates the successful nature of the interaction. In contrast, no restriction events were directly observed on 15 mM Ni^{2+}surfaces throughout the duration of the study.

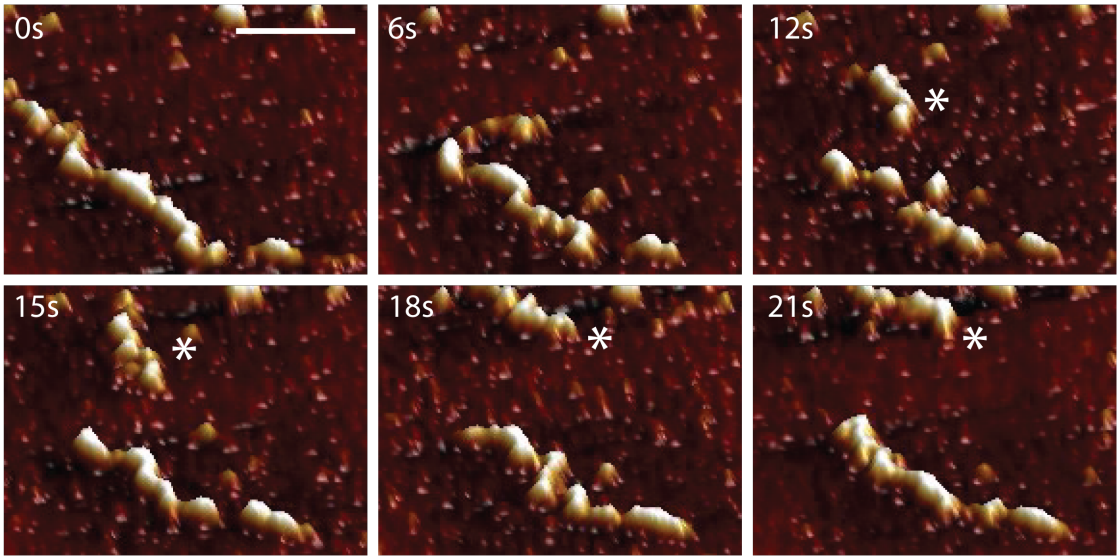

Figure 7.10: ***Sequential HS-AFM images depicting the successful restriction of a DNA molecule immobilised on a 5 mM Ni^{2+}mica surface.*** *The DNA is cleaved (between 6 s and 12 s) and a short DNA fragment (indicated by a white star) is seen moving towards the edge of the observed area. The two resultant DNA molecules are seen to move independent of one another indicating a complete cut. Scale bar is 200 nm. Z scale is 6 nm.*

7.2.8 Application to the Observation of RecA Polymerisation on Double Stranded DNA

Having proven the sample preparation methodology on an exemplary system with great success, attentions turned back to the RecA system which is the core of this project. In the first instance, attempts were made to observe the pre-synaptic phase of the RecA homologous recombination – the formation of a NPF (see section 3.2.3). It must be noted that this is a considerably challenging event to attempt to orchestrate in a surface confined regime, when the structure of the final RecA polymer is considered (section 3.2.2).

Where 3.5 kbp DNA is immobilised on a 10 mM Ni^{2+}- mica substrate and RecA monomers are

slowly introduced, evidence of nucleation becomes evident (figure 7.11, white arrow & movie 7.5). Note that RecA is pre-incubated with ATPγS, and diluted in the imaging buffer prior to addition. A variety of RecA ratios were explored, from 0.1:1 - 10:1 (RecA:DNA), with a 1:1 ratio found to confer the best results. Ratios below 1:1 were found to lack the concentrations necessary to nucleate on the surface bound DNA, typically no interactions were observed over the course of 2 hours – after which surface crystallisation fouled the experiment. In contrast, ratios above this primarily resulted in the accumulation of RecA monomers on the substrate, making the distinction of specific interactions impossible.

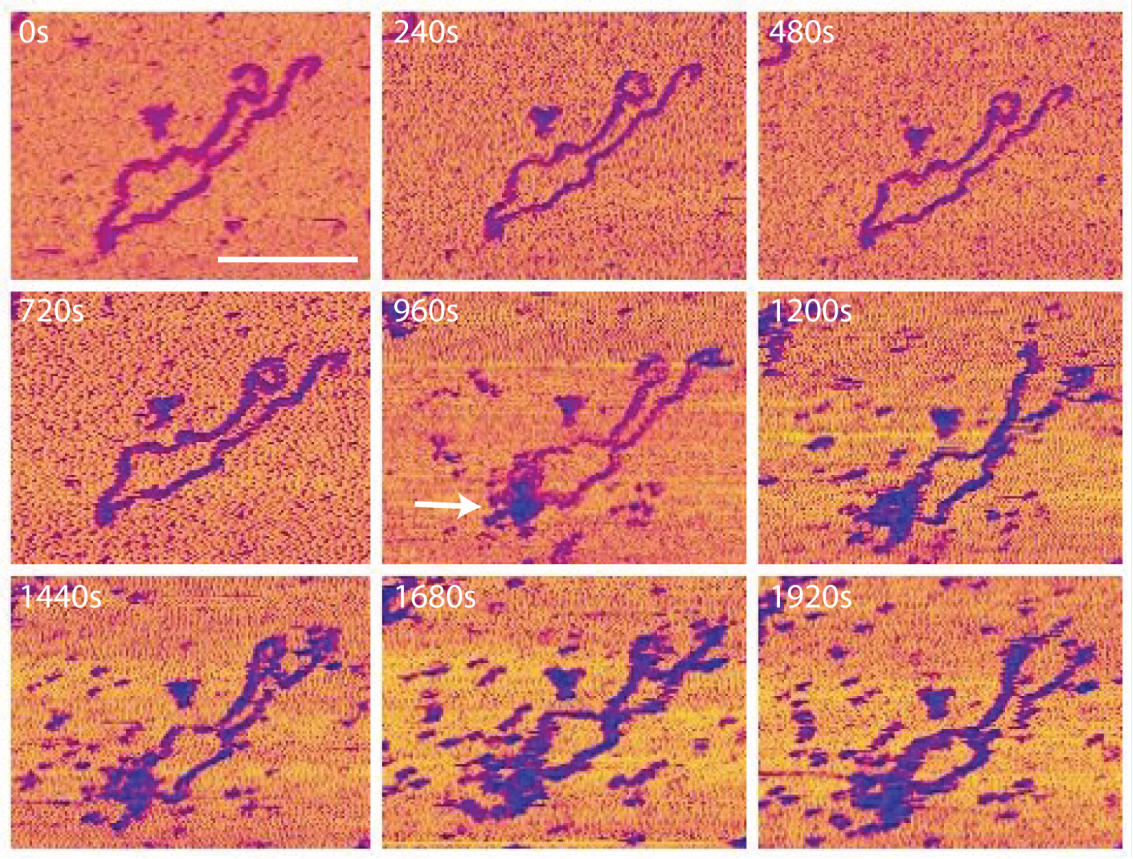

Figure 7.11: ***HS-AFM observation of RecA nucleation.*** *A series of HS-AFM images depicting the nucleation of RecA monomers on 3.5 kbp dsDNA. RecA monomers are seen to cluster around a loop in the DNA template (white arrow) indicative of nucleation. Scale bar = 250 nm. Z scale = 8 nm.*

As such, these experiments were highly susceptible to protein fouling as a consequence of the large excess of protein required to provide the critical nucleation concentration. This nucleation phase is known to be rate limiting (section 3.2.3) and hence is the critical to observing full polymerisation.

Evidence for the formation of nucleation clusters at specific points on the bound DNA molecules can clearly be seen in (figure 7.11 & movie 7.5). However, throughout the experiments, there

was no suggestions of an extension phase observed. This could be a consequence of the degrees of freedom of the DNA molecule still remaining limiting, where nucleation clusters are seen to form at the most flexible regions. Alternatively, it is likely that under the concentrations where nucleation events were observed, the critical mass is not reached for extension to proceed. It maybe considered that the higher concentrations explored may have circumvented this, however due to surface fouling any events occurring remained unresolved.

Despite being important for an understanding of the overall homologous recombination mechanism, such experiments of the pre-synaptic phase may be considered peripheral to the scope of this thesis. With this in mind and given the evident difficulty of these experiments, it was concluded that they would be better pursued in future investigations once solid investigations of the RecA homology searching mechanism had been concluded. Nonetheless, these initial investigations suggest that RecA proteins are able to undertake interactions with dsDNA sufficient to give confidence in this experimental approach.

7.2.9 RecA Nucleoprotein Filament Searching Behaviours Observed

The observation of RecA NPF interactions offers several advantages over the previous experiments of RecA polymerisation. Primarily, the concentrations required are drastically lower and the stable NPF complex is less susceptible to fouling the surface.

In the first instance, the surface diffusion of 60 nt NPFs was assessed. NPFs were formed at a concentration of 1 μM as previously described (methods 5.2.6) and deposited in Tris - Mg buffer (methods 5.1.3) on 10 mM Ni^{2+}- mica and imaged without further incubation. From movie 7.6 NPFs can clearly be seen to diffuse freely across the surface at $22°C$ and was considered sufficient for the intended experiments.

The DNA:NPF dynamic experiments were setup as discussed for the previous observations with EcoRI and monomeric RecA. 3.5 kbp was deposited on 5 and 10 mM Ni^{2+}- mica, as discussed previously and stable imaging established prior to the introduction of 60 nt NPFs. Again a variety of interaction ratios were explored from 3:1 - 0.1:1 (NPF:DNA), with typically ratios of 0.1:1 and 0.5:1 offering the best results and higher ratios characterised by surface fouling.

Interestingly, a host of observations began to shed light on NPF homology searching interactions, with multiple different events occurring on a single DNA molecule throughout a 10 minute observation window. Figure 7.12 (& movie 7.7) depicts the interaction of a single NPF (white arrow) and a loop of the 3.5 kbp DNA template. Additional NPFs can clearly been seen

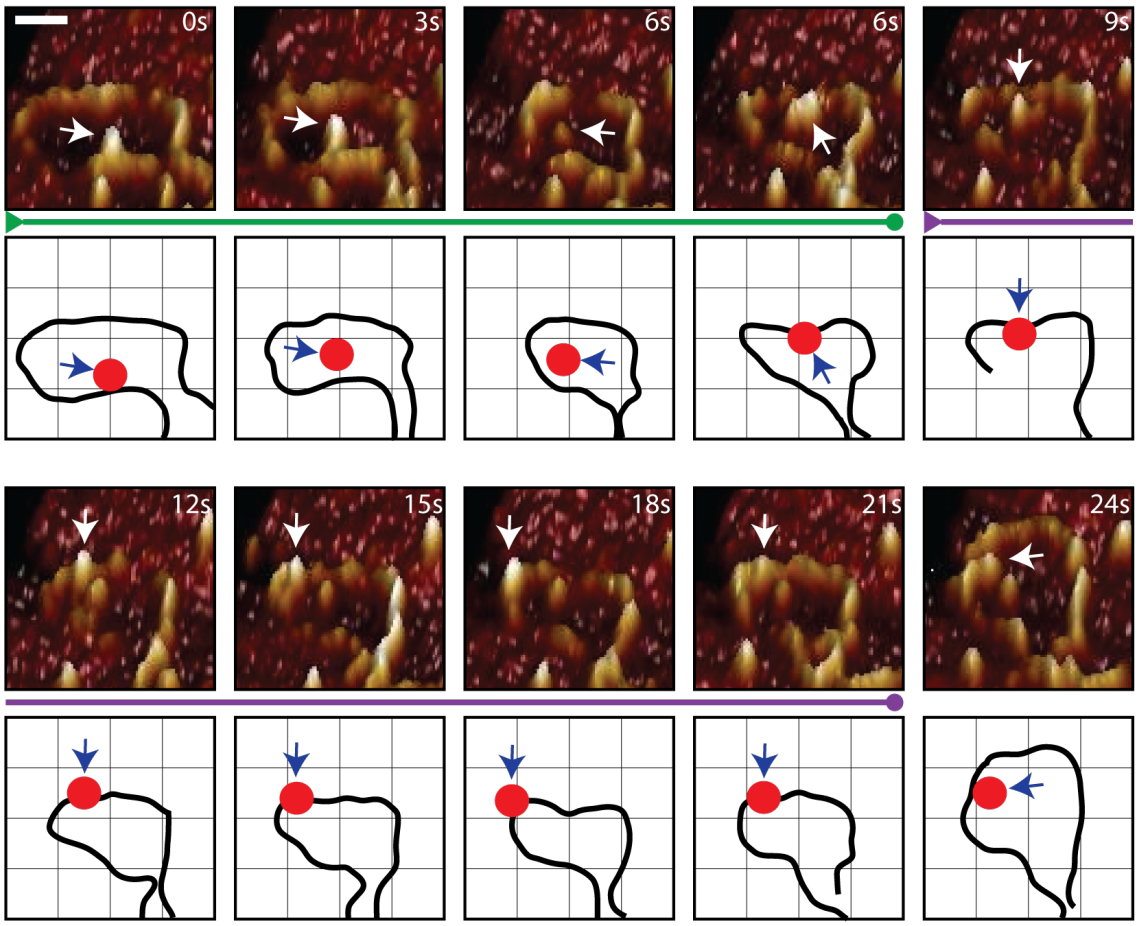

Figure 7.12: ***HS-AFM observation of the homology searching interactions of RecA NPF with dsDNA*** *A 60 nt RecA NPF (white arrows) is observed to interact with a loop in the 3.5 kbp DNA template through a combination of random interaction (green time line) and facilitated diffusion (purple time line) over a short distance. Shown as a series of HS-AFM images and vector plots. Scale bar = 25 nm. Z scale = 8 nm.*

surrounding the template DNA. For clarity, all the molecules observed were converted to vector representations (see section 7.2.3) and plotted such that their relative motions could be tracked (figure 7.12 & movie 7.7).

From these observations NPFs are seen to diffuse across the surface, or appear directly from solution (see figure 7.13 & movie 7.8), forming random initial contacts with the partially immobilised DNA. Interestingly, these contacts are observed to be transient in some cases, "hopping" to an additional site further down the template (figure 7.12, green timeline) which is suggested to be the primary mechanism of interaction (section 3.2.3).[192,193,194,188]

However, in the work presented here clear evidence is seen for the existence of a one dimensional facilitated diffusion (sliding) of NPFs along dsDNA (figure 7.12, purple timeline & movie 7.7) – albeit covering only short sequence distances of up to approximately 50 bp. This is found to

fall short of the work by Ragunathan et al[15] (see section 3.2.3) which suggested slides of up to 300 bp in length - approximately six times the distances estimated here. This discrepancy is likely a result of the experimental setup, where confinement at a solid-liquid interface in this work inevitably retards the interaction. Nonetheless, the findings presented here still provide direct evidence supporting the existence of facilitated diffusion within the RecA homology searching mechanism.

Interestingly, as was noted in the previous chapter (section 6.2.6), RecA NPFs are seen to form clusters in the vicinity of the DNA molecule. It could be argued that, this is merely an effect of all the NPFs individually targeting the same DNA molecule and hence local concentration enhancement is a result of the presence of the nucleic acid itself. However, from figure 7.13 (and movie 7.8) a cluster of NPFs are clearly observed to arrive simultaneously from solution suggesting a previous cooperation in the absence of the DNA molecule. The instantaneous appearance of this cluster and the subsequent diffusion of individual NPFs away from the cluster, is strikingly similar in its appearance to those observed in the static experiments of the previous chapter (section 6.2.6). Furthermore, from these observations it is evident that these clusters are formed from weak associations of NPFs and are not indicative of detrimental aggregation (figure 7.13). As such, this data further supports the existence of a massively parallel homology search orchestrated by cooperation between individual RecA nucleoprotein filaments.

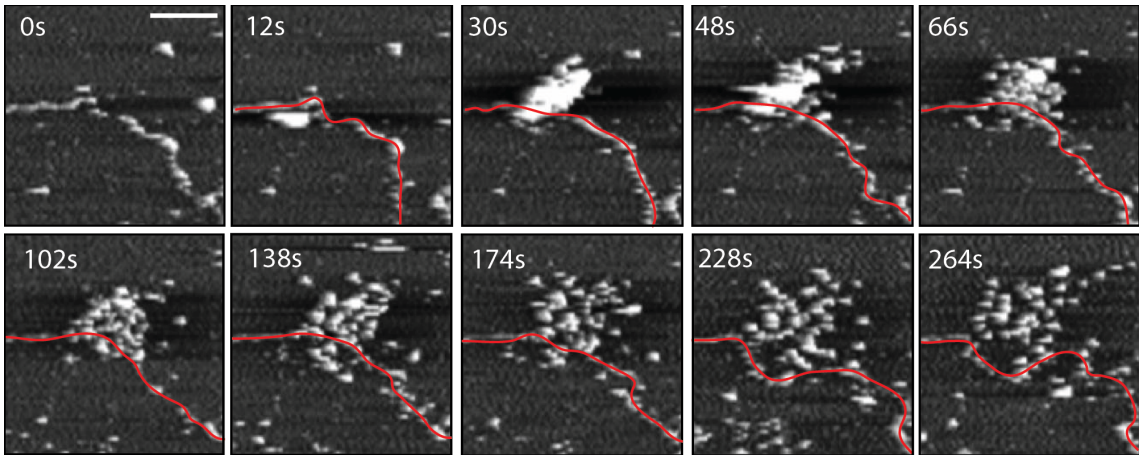

Figure 7.13: ***Real time observation of nucleoprotein filament clustering*** *A series of HS-AFM depicting the appearance and interaction of a cluster of NPF molecules with 3.5 kbp DNA. Scale bar = 100 nm. Z scale = 8 nm.*

7.3 Conclusions

When attempting to observe biological processes in vitro with the HS-AFM, it is essential to ensure that the measurement only minimally perturbs the processes and interactions of interest. For the investigation of nucleoprotein interactions, the subject DNA molecules must be associated with the surface strongly enough to enable stable imaging over relevant time-scales, while at the same time providing enough translational freedom to allow the proteins to interact with the DNA in a way that is as undisturbed as possible. A number of different approaches have been reported in the literature to control the surface association of DNA molecules.[326,327,328,329,330,331] However, none of the existing methods offer the ability to tune the surface interaction and thus the translational freedom of the DNA appropriately to enable a large range of diverse investigations within the additional constraints of maintaining suitable reduced physiological conditions.

The work in this chapter has demonstrated a novel approach to tune the surface interaction of DNA which fulfils these requirements. Both stable imaging and reaction conditions are orchestrated simultaneously without the need to perturb the system throughout the reaction time-course, through cyclic buffer exchanges or electric potentials.

In order to do this, an analytical method for quantifying the mobility of DNA under different surface association conditions was developed, enabling the interaction of the DNA with the mica surface to be tuned from a tight surface-association to highly mobile one. This approach, therefore, provided a highly flexible method to establish conditions which enable the investigation of a large range of nucleoprotein interactions, as demonstrated here on EcoRI and RecA. Further to this, the presented analytical approach enables the analysis of the impact of various conditions to the subject molecules, including the interaction of the AFM tip and the existence of specific surface pinning points.

These approaches were successfully applied to the observation of RecA dynamics, in particular the homology searching behaviours of NPFs on dsDNA. These direct observations, provide strong evidence for the existence of a facilitated diffusion along the DNA template, the existence of which is debated by others.[192,193,194,188] Moreover, the prevalence of cooperative action is demonstrated through the observation of NPF clusters, providing further evidence in support of the observations made in the previous chapter (section 6.2.6).

However, despite these successes, due to the inherent mobility of both molecules (DNA and NPF) orchestrated by this sample preparation method, it remains difficult to reconcile specific quantities

of the interactions. In particular, it is difficult to infer any directionality, without the additional inclusion of a polarity marker to the DNA template. Furthermore, quantifying the distance and hence velocities of NPF sliding is complicated by the lack of fixed references from which to measure.

In order to overcome these limitations, the next chapter explores the use of DNA origami as a nano-scale experimental platform to both anchor and provide positional references for the template DNA molecule, enabling the above quantities to be examined.

This page is intentionally left blank.

Chapter 8

Observations of RecA Interactions within DNA Origami Support Structures

8.1 Introduction

The experiments conducted in the previous chapter directly observed a host of RecA interactions with DNA. These observations provided direct evidence for the occurrence of NPF facilitated diffusion along dsDNA as well as clustering behaviour. However, despite these successes, the extraction of detailed quantities regarding the interaction between NPFs and DNA species remained difficult with the methods applied therein.

Consider the demonstrated example of the *in situ* DNA restriction by EcoRI. Here, successful interactions are evidenced by a large scale manipulation of the DNA, in this case its division into two fragments, an unambiguous state change. In contrast, quantifying processive movements,[127] structural re-arrangement of nucleic acid,[345] mechanical dependencies of nucleoprotein interactions,[346,122] or discrete stages of motor activities of enzymes[231] requires detection of subtle changes. As discussed, this insight is crucial in understanding the operational constraints of these nucleoprotein interactions when considering harnessing them for bionanotechnological application.

The sample preparation methodology discussed in the previous chapter provided sufficient translational mobility for both nucleic acid and protein species simultaneously in an attempt

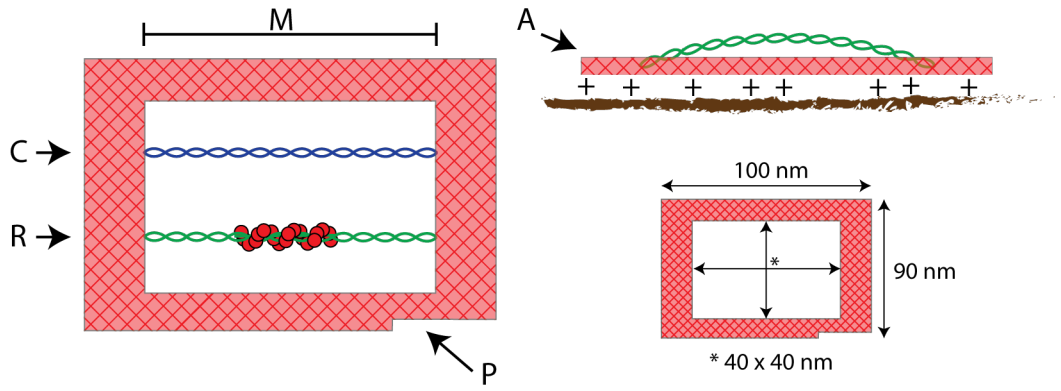

Figure 8.1: ***An overview of the DNA origami frame method.*** *The structure and various functions of a DNA origami frame are depicted. The frame acts as a measurement standard (M) and an orientational reference (P), whilst simultaneously anchoring the reaction to the substrate (A). Reactions can be observed on internally incorporated DNA strands (R) alongside simultaneous controls (C). The dimensions of the frame used in this study are given.*

to emulate the 3D interaction environment within a confined 2D plane. However, it became evident that in a regime where both interacting species are able to move freely about one another, quantifying the motion of one relative to the other is very difficult.

Hence, the spatial position of the DNA must ideally remain consistent and yet accessible to the NPFs. This again presents a juxtaposition of requirements within the experimental design. One might consider that a way to meet these requirements would be to tether the DNA termini at fixed distances from one another, such that the DNA molecule lies in a predictable manner between two loci.

Over the years, several approaches for tethering DNA to surfaces have been demonstrated - typically for the formation of ssDNA arrays for sequencing applications.[347] However, the substrate roughness or chemistries of these approaches are found to be unsuitable for application in this work. Further, these methods result in the random deposition of molecules making the spacing of two defined loci impossible to achieve reliably.

Alternatively, an elegant solution demonstrated by Sugiyama and Endo et al[128,236] is to use DNA origami (section 3.1.3) to define the spatial separation of the termini of a linear dsDNA strand. Here, M13 viral ssDNA is woven to create a frame structure containing a central window within which reactions can be hosted (figure 8.1), such as site specific recombination of DNA by Cre recombinase.[236] Due to the increased surface area, the origami structure stably binds to mica through the interstitial Mg^{2+}interactions - utilised in the previous chapters - without the need

for the coordinated bonding of Ni^{2+}. This enables the central dsDNA strand to remain free for interaction yet presented across a predictable location (figure 8.1). Furthermore, the highly structured architecture of DNA origami can be utilised as a measurement standard against which the position and orientation of the NPF can be quantified (figure 8.1).

In this chapter, the DNA origami frame methodology is adopted to circumnavigate the limitations in quantifying position and orientation of the subject molecules in previous experimental system (chapter 7). Here, the NPF homology search is conducted using DNA origami as a measurement standard in an attempt to quantify the interaction orientation and rate. In the first instance, a design was re-created from the literature in order to conduct initial investigations. Later adaptation of this frame design to increase the size of the resident dsDNA molecules to enable further homology search observations and potentially provide a platform for direct observation of full RecA mediated strand exchange was undertaken in collaboration with Masayuki Endo (Sugiymama Group, Kyoto University, Japan).

8.2 Experimental

8.2.1 Generating the Requisite DNA Origami Frame

For the purpose of observing the NPF homology searching mechanism a DNA frame that could host two dsDNA within its central window was required, such that an entirely heterologous sequence may be included as a control throughout the observations. A suitable design was re-created from Yamaoto et al[128,122] which is referred to as DF2S throughout this text. This design included a central window of 40 x 40 nm and a clear polarity marker enabling the orientation of the frame and its hosted molecules to be determined (figure 8.1).

From the published schematic diagram, the rough layout of the scaffold and staple components were known along with the structural dimensions.[122] The design was recreated in cadnano (section 5.2.9),[314] using standard origami design rules on a square lattice in 8 bp sequence fractals. This resulted in the standard placing of staple crossovers 16 b apart - or every one and a half helical turns. Following the specified dimensions of 100 nm by 90 nm, the design was produced from a weave of M13mp18 ssDNA with 224 oligonucleotide staples, resulting in a width of 288 bp and a height of 32 duplexes (figure 8.1). The cadnano schematic can be seen in the appendix (figure A.1).

The 40 x 40 nm central window was created through the omission of 128 bp by 16 duplexes at

the heart of the design (figure 8.1). Furthermore, 16 nt ssDNA anchor points were included to enable insertion of two dsDNA in parallel within the window (figure 8.1). These anchor points projected into the window by 16 nt and were therefore 32 nt long in total (see appendix, figure A.1). From the published design, the sequence of these ssDNA anchor regions was known and hence was used as a reference to populate the sequence throughout the whole origami using a M13mp18 bacteriophage ssDNA template.

8.2.2 Preparation of the DNA Origami Frame

Initial preparation of the DNA origami frame was conducted at a variety of staple:scaffold ratios to establish the appropriate folding conditions for high product yield. Staple mixes were formed by pooling the 224 staple oligonucleotides and introducing them to M13mp18 in Tris-Mg buffer (methods 5.1.3) at a variety of excesses to assess the formation of the DF2S frames. The mixtures were heated to 95°C and folded by steadily cooling back to room temperature using a thermocycler, reducing the temperature by 1°C min^{-1}.

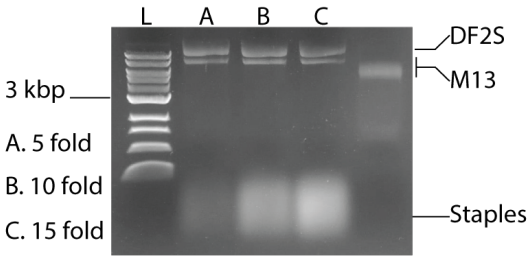

Figure 8.2: ***Assessment of DF2S DNA origami frame folding.*** *A 1% agarose gel demonstrating the successful folding of the DF2S DNA origami frame with a 5 (A), 10 (B) and 15 (C) fold excess of staples. The folding efficiency is found to be saturated already at 5 fold. L = 1 kbp ladder.*

The successfully folded DNA origami frame is distinctly resolved at the top of each lane when separated in a 0.5% agarose gel (figure 8.2, lanes A - C). As the folded product is considerably more rigid and structured than its constituents, it is severely retarded by the gel matrix. As a consequence, the folded origami structure is seen to be distinct from the constituent M13mp18 ssDNA template run as a control in the last lane. Interestingly, the DNA origami is represented by two bands which could indicate the formation of additional structures, such as dimerisation or rolling of the 2D sheets into filamentous geometries. Evidence for both of these states were observed in subsequent AFM experiments. Regardless, the DNA frame was found to fold with an estimated 66% efficiency across all three staple mixtures examined. Further, it was evident that the folding reaction already had a saturated product yield with a 5 fold excess of staples, more than sufficient for the subsequent work.

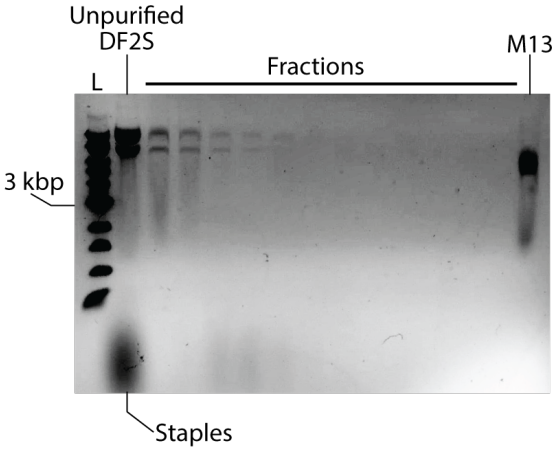

Figure 8.3: ***Removal of excess staples from the DF2S DNA origami frame.*** *A 1% Agarose gel depicting fractions of DF2S purified with S400 size exclusion columns. The staples are efficiently removed yielding the majority of pure product in the first minute. Each fraction is produced by spinning the column at 750 rcf for 30 s in a microcentrifuge. L = 1 kbp ladder.*

Furthermore, the excess staples are clearly visible at the bottom of the gel in figure 8.2 (lanes A - C), increasing in intensity with excess concentration. As these would contaminate any subsequent experiments with RecA, they were removed with high efficiency using a sephacryl S400 size exclusion chromatography matrix (methods 5.2.9), utilised here in a spin column format due to the small quantities of product utilised in these studies. Here, the highly structured DNA origami is too large to diffuse through the porous matrix and is hence excluded to the reduced volume between the gel beads and as such, is eluted first (see methods 5.2.9). The excess staple oligonucleotides – with an average length of 32 nt – interact with the pores of the matrix and hence diffuse through a larger volume. Figure 8.3 shows the efficient removal of excess oligonucleotide staples from the folded DNA origami frame, with the product eluted primarily in the first 2 fractions – where each fraction constitutes centrifugation at 750 rcf for 1 minute intervals. The staples remained within the column through a subsequent 10 fractions. It is important to note that this purification approach was not sufficient to separate the partially folded M13mp18 from the product due to the similarity in MW. Nonetheless, this method was found to be sufficient for all downstream experiments.

8.2.3 Design of the Internal dsDNA Strands

In order to conduct NPF homology searching observations, two pairs of dsDNA central strands were designed to be incorporated within the DNA origami frame. Each pair contained a reaction and a control strand (figure 8.1), where the latter was completely heterologous to the NPF used in the experiment. The reaction strand was always inserted closest to the polarity marker in the DNA frame for unambiguous reference.

Within the 128 bp wide central window, a 64 bp dsDNA – representing 6 full helical turns of B-form DNA – could be accommodated between each the pair of 32 nt anchor points (figure 8.4).

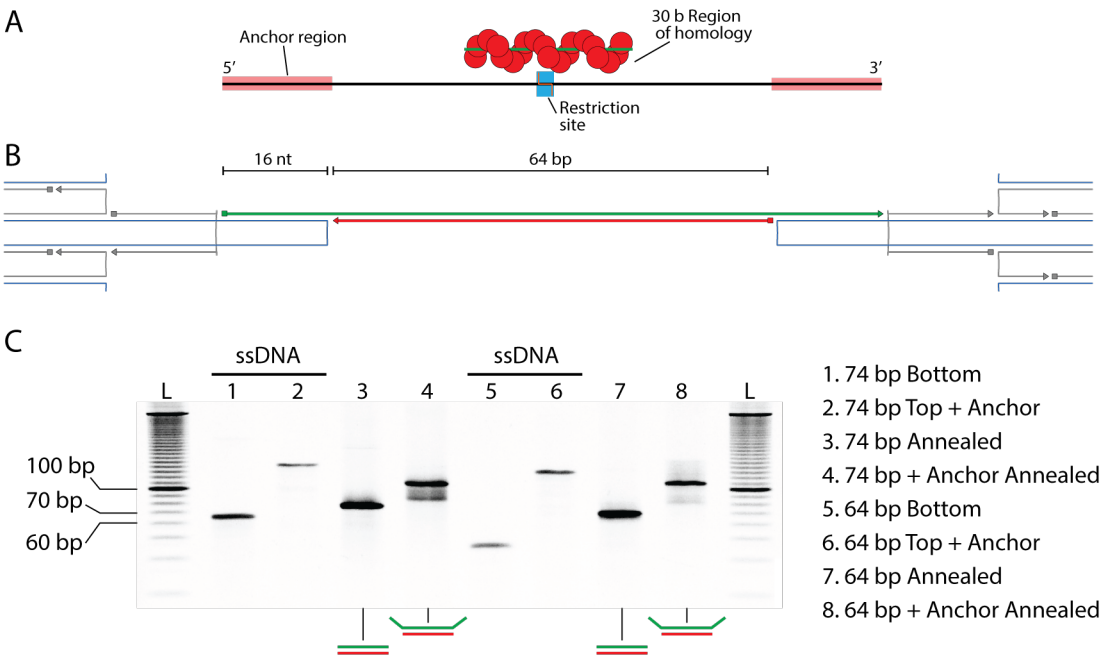

Figure 8.4: ***The design and formation of the DF2S central strands.*** ***A)*** *A schematic diagram depicting functional layout of the central reaction dsDNA strand, including anchor regions, patterning region and the location of a restriction site.* ***B)*** *A schematic diagram depicting the composition and binding scheme of the central strands, within the context of the DF2S binding loops.* ***C)*** *The formation of the 64 bp and 74 bp versions of the central strands is shown in a 15% polyacrylamide gel. The individual ssDNA components and the final annealed products are shown demonstrating high yields. The inclusion of the anchor sequences results in the multiple bands when annealed due to the presence of secondary structure in the ssDNA overhangs. L = 10 bp ladder.*

As degrees of freedom of the DNA substrate is shown to be critical to the interaction,[11] a second version was produced where the flexibility of the dsDNA was increased with an additional helical turn, giving a total length of 74 bp.

The central dsDNA strands were designed to contain terminal ssDNA overhangs complimentary to the origami frame anchor points. 30 nt NPFs were generated from ssDNA homologous to the centre of the upper strand within the dsDNA region, across a EcoRI restriction site (figure 8.4 A). The central strands were formed from two complimentary oligonucleotides, one of which included 16 nt ssDNA overhangs at either termini which were complimentary to the origami frame anchor sequences (figure 8.4). Hence, the 64 bp central dsDNA was formed by annealing 64 nt and 96 nt ssDNA. Arbitrary sequences were extracted from λ bacteriophage DNA and modified to include a central EcoRI restriction site (figure 8.4 A). Figure 8.4 C, shows the high yield formation of both central strand length variants when annealed in 10 mM Mg^{2+}containing buffer (Tris-Mg buffer,

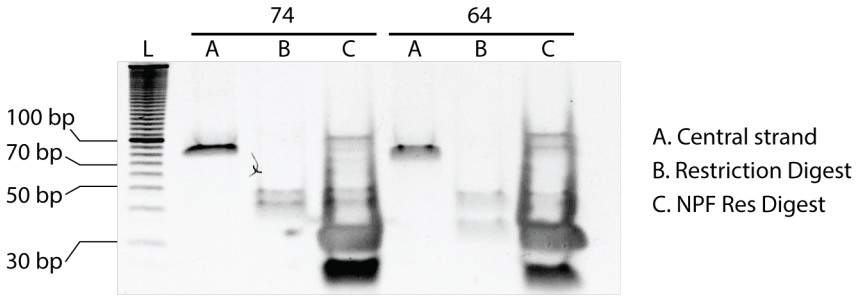

Figure 8.5: ***Restriction assay assessment of DF2S central strand patterning.*** *A 15% polyacrylamide gel demonstrating successful patterning and protection of the 74 bp and 64 bp templates by homologous 30 nt NPFs. Note due to the similar size of the patterning oligonucleotide and restriction fragments the resolution is diminished. L = 10 bp Ladder.*

methods 5.1.3). It is important to note, that the central strands resolve higher in the gel – with respect to the 10 bp ladder – than expected and give rise to multiple bands, due to the presence of ssDNA overhangs which are capable of forming a variety of secondary structures (figure 8.4 C, lanes 4 & 8).

8.2.4 Central Strand NPF Patterning Restriction Assay

Prior to incorporation of the central strands into the DNA origami frame, the patterning efficiency was established as previous, using a restriction assay (see methods 5.2.7). From figure 8.5, protection of both the 74 bp and 64 bp central strands from restriction by EcoRI is afforded at a 10:1 (NPF:DNA) patterning ratio (lanes C). In contrast full digestion of the dsDNA substrates is seen in the absence of NPF (figure 8.5, lanes B). It is important to note that due to the short restriction fragment lengths and their comparable size to that of the NPF oligonucleotide – and the further oligonucleotide included to bind excess RecA – which are included at excess, that accurate densitometry assessment is difficult. That being said, the interaction efficiency was estimated to be approximately 30% which was considered sufficient for the requirements of this chapter.

8.2.5 Incorporation of the Central dsDNA into the DNA Frame

Incorporation of the central reaction strands was conducted post-formation with the central strands provided in excess, where a 2:1 ratio (central:frame) was found sufficient. Incorporation was conducted by heating the reaction mixture to 40°C and cooling at -0.5°C min^{-1} to room temperature in a thermocycler. Figure 8.6 depicts the removal of excess central strands and staple oligonucleotides in the same step, as described previously (section 8.2.2). Where purification is

undertaken at 750 rcf (relative centrifugal force) with fractions of 30 second spin intervals, the central strands are seen to already appear in the second fraction. No appreciable improvements were gained by reducing the spin speed to 500 x g and hence these losses were considered tolerable. It is important to note that in this spin column format, the fraction limits are considered relative and improvements could be made by undertaking a more rigorous fast protein liquid chromatography (FPLC) setup. However, this would require considerably larger amounts of product for purification and was deemed unnecessary.

8.2.6 AFM Assessment of DNA Frames

As there is no way to identify the inclusion of the central strands from gel electrophoresis, AFM was used to characterise the purified DNA frames. Typically, 15 ng was deposited on mica and imaged in the folding buffer - which contained 10 mM Mg^{2+}, sufficient for binding of the origami (see methods 5.2.8). Figure 8.7 B and C, shows representative images of DNA frames before and after incorporation of the central dsDNA strands, respectively. From these images the inclusion of the central dsDNA strands was estimated to be approximately 90% efficient at the 2:1 ratio used here. This was found to be in good agreement with the original published literature.[128,122]

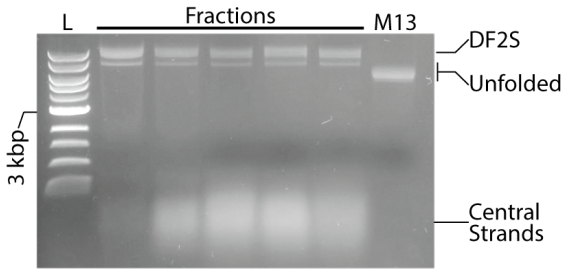

Figure 8.6: ***Removal of excess central strands from the DF2S frame following incorporation.*** *A 1% agarose gel depicting fractions of DF2S following incorporation of the central strands, purified with S400 size exclusion columns. The central strands are removed in the first fraction, with a low but sufficient yield of DF2S. Each fraction is produced by spinning the column at 750 rcf for 30 s in a microcentrifuge. L = 1 kbp ladder.*

Interestingly, defects in the frames are clearly evident, including missing staples or partial tears in the weave (figure 8.7 D). Where staples appear to be missing, this is likely a consequence of the folding pathway and does not appear to be harmful to the overall structure, indicating a degree of redundancy. However, the appearance of tears - typically observed running parallel to the helical axis of the constituent DNA - are more likely evidence of detrimental interaction with the HS-AFM probe. This highlights that, despite being reasonably robust structures, DNA origami frames are susceptible to damage during HS-AFM imaging, which was critical to consider in later experiments.

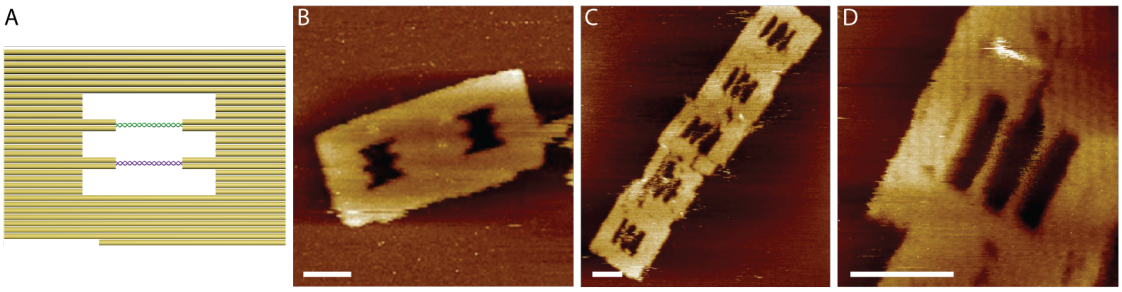

Figure 8.7: ***AFM assessment of DF2S and incorporated strands.*** ***A)*** *A schematic representation of the DF2S frame including central strands.* ***B)*** *A representative AFM image depicting empty DF2S frames.* ***C & D)*** *A set of representative images depicting DF2S frames including central strands. Scale bars = 40 nm. Z scale = 3 nm.*

Interestingly, the majority of DNA frames were observed as dimers or longer chains, contrary to the mono-dispersion expected. This is found to be a consequence of base–stacking interactions between the termini of the helices in neighbouring DNA origami frames and as a result the frames form polymers along this axis.[348] This phenomena can be rectified through the inclusion of poly T regions on the terminal staple sequences which extend out into solution.[348] However, a redesign to include this was deemed unnecessary as despite being unexpected, the appearance of these chains helped to group the frames together in a systematic way, enabling multiple to be observed simultaneously.

8.2.7 Assessment of the Mg^{2+}Dependence of the DNA Origami Frames

In order to further assess the utility of the DNA origami frames for observing RecA interactions with DNA, their stability in low Mg^{2+}conditions was examined. This is important as Mg^{2+}concentration is the key to controlling RecA specificity. In particular, concentrations below 2 mM enable targeting of ssDNA, whereas concentrations above this allow dsDNA binding to be specified. DNA frames were folded and the central strands incorporated using Tris-Mg buffer (methods 5.1.3) – containing 10 mM Mg^{2+}, as discussed. These DNA frames were subsequently exchanged into Tris-Mg buffers containing a variety of Mg^{2+}concentrations – from 0 – 10 mM – during the purification step by pre-equilibrating the size exclusion matrix prior to introduction of the sample.

In figure 8.8, the effect of reducing Mg^{2+}concentrations on the stability of the DNA frames is demonstrated. The DNA origami structure is found to progressively lose structural integrity with decreasing Mg^{2+}concentration, associated with a reducing resistance to interaction with the HS-

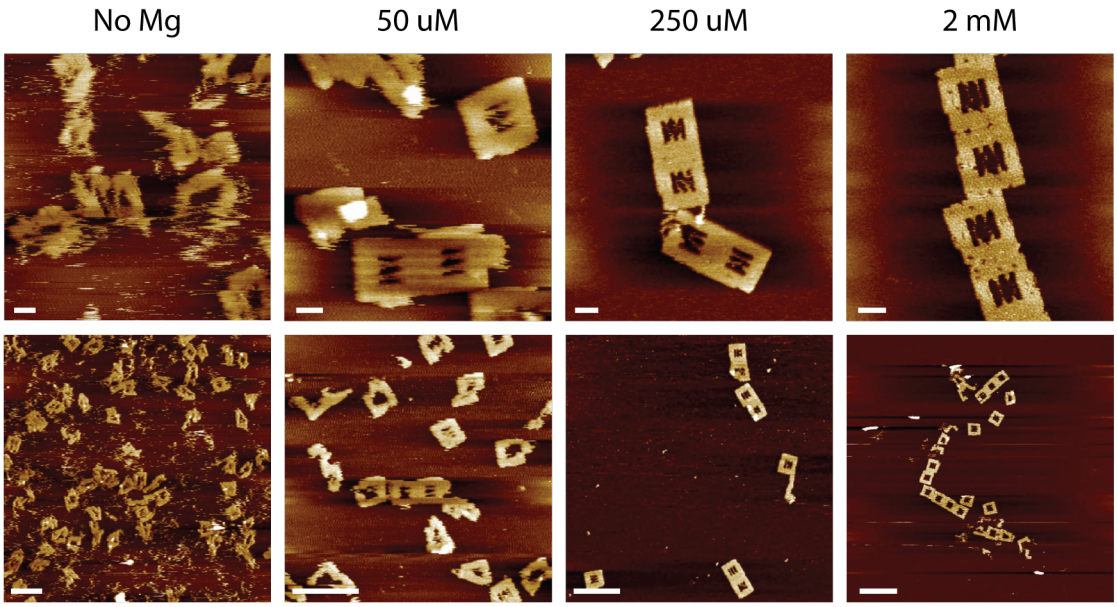

Figure 8.8: ***Assessment of the stability of DF2S frames under different Mg^{2+} concentrations.*** *Representative AFM images of DF2S frames formed at 10 mM Mg^{2+} and exchanged into a variety of Mg^{2+} concentrations. Frames are seen to be completely stable down to 250 μ M. Scale bars, top = 40 nm, bottom = 250 nm. Z scales = 4 nm.*

AFM probe (figure 8.8). DNA frames are found to remain reasonably stable down to 250 μM Mg^{2+}, sufficient for consistent imaging over 30 minutes. However, note the increased appearance of multiple defects at concentrations of 2 mM Mg^{2+}. Below 250 μM Mg^{2+}, it became difficult to reliably observe any structure for more than a few consecutive images. This was found to be particularly acute where Mg^{2+} was removed entirely, as the DNA origami was seen to unfold at the surface (figure 8.8, no Mg) - characterised by the inability to resolve the DNA as it is moved around by the HS-AFM probe. As such, the DNA frames are found to be robust to a variety of Mg^{2+} concentrations above 250 μM – and certainly below the 2 mM threshold desired – enabling experimental versatility.

8.2.8 Nucleoprotein Filament Homology Searching within DF2S

Having established the protocols for producing DNA frames at sufficient yield - despite appreciable losses through purification - experiments to observe NPF homology searching were conducted. 30 nt NPFs formed as described previously (section 5.2.6) were introduced to DNA frames on a mica surface at the same ratios as utilised in the previous chapter (see section 7.2.9), typically a ratio of 0.1:1 (NPF:DNA) was used.

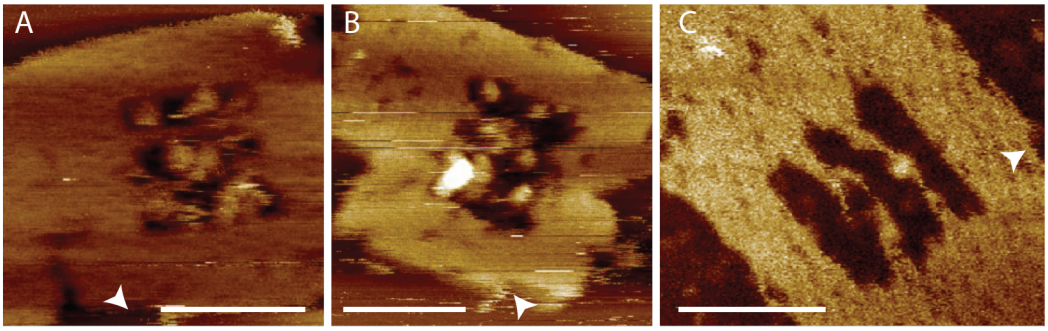

Figure 8.9: ***Interactions of NPFs within DF2S conducted on mica without continuous HS-AFM imaging.*** *Representative AFM images depicting the non-observed patterning of 30 nt NPFs inside the DF2S frame.* ***A & B)*** *Clustering behaviours are difficult to overcome even at low concentrations.* ***C)*** *Observation of correct patterning of singular NPF. Arrows indicated the position of the polarity marker. Scale bars = 40 nm. Z scale = 6 nm.*

Figure 8.9 shows representative images of samples incubated for 30 minutes at $22°C$ on a mica surface without continuous observation. NPFs were seen to cluster within the DNA frame central window – even at these low stochiometries – in-line with the cooperative interactions noted in the previous chapters (see sections 6.2.6 & 7.2.9). Where these NPF clusters are observed, these DNA frames are ignored as it is impossible to discern any useful information (figure 8.9 A & B).

Interestingly, singularly interacting NPFs were typically found on the periphery of these larger clusters, suggesting a degree of sacrificial DNA frames must be tolerated to overcome the NPF cooperative behaviours. Figure 8.9 C depicts a representative HS-AFM image of a completed homology search. The NPF can clearly be seen bound to the central region of homology on the reaction strand (for reference see figure 8.1), with respect to the polarity marker (white arrow).

These observations further validate that NPFs do not require a free DNA termini to mediate an interaction. However, one must consider that as the central strands are bound into the frame using only ssDNA overhangs, that a backbone nick exists at either termini. Hence, one can not rule out that the interaction of NPFs may be enhanced at these lesions. This is important where future intentions are to pattern NPFs within complex DNA structures – hence the inclusion of nicks may be necessary to enable this to occur.

Where these interactions are observed in real-time with HS-AFM, clear evidence of NPF homology searching behaviours are evidenced. NPFs are seen to move freely within the central cavity of the DNA frame, tracking back and forth along the DNA strands therein (figure 8.10 & movie 8.1). This is found to be consistent with the sliding behaviours evidenced in the previous chapter (section 7.2.9).

Moreover, these motions are apparently irrespective of the orientation of the frame with respect to the scanning probe and hence can be considered largely distinct from its influence. This can clearly be seen in figure 8.10 A, where two DNA frames are shown set at 90° from one another, being broadly perpendicular and parallel with the fast scan (X) axis, left and right, respectively. Where the centre of mass of the NPFs is plotted with respect to the trajectory of the DNA strand – throughout successive images, similar motions are observed for both orientations (figure 8.10 B). Note that for clarity, only the DNA strand upon which the NPFs are conducting a homology search is included in the plot and that as it is the centre of mass that is being tracked here, the data points plotted appear to be distinct from the DNA, which may be misleading.

8.2.9 Distinction of Two NPF–DNA Interactions

Interestingly, as described in section 8.1, the DNA frame can be used as a positional reference. Hence the trajectory of NPF along the subject DNA molecule can be tracked with respect to the distance from the polarity marker. This allows the identification of discrete steps in the homology searching interaction to be discerned, a representative example is given in figure 8.10 C. Here, existence of two distinct motions can be discerned, characterised as “hops” - where the NPF jumps from one location to another in a random fashion - and “slides” - where a progressive uni-directional movement is observed across short distances.

Figure 8.10 suggests that the NPF generally hops from one location to another, typically covering distances of 38 bp, presumably undertaking transient homology probing. However, short regions of progressive sliding are clearly observed, with a maximum of 25 bp seen here. On average, slides are observed over distances of 10 bp – which is in-line with the short distances observed in the previous chapter on linear templates (see section 7.2.9). Examples of these distinct steps can clearly be extracted from the raw HS-AFM data, such as the representative events in figure 8.10 D and E, depicting sliding and hopping, respectively.

These observations are found to be in-line with previous discussions of a two stage homology probing process. This is interesting, as recent simulations have suggested that the existence of metastable probing complexes - existing as an intermediate between the pre-synaptic and stable synaptic phases - could account for the speed and stability paradox of the RecA homology search.[160,349,350,351] These studies suggest that continually transient probing, where the incoming DNA primarily exists in the binding site II, could result in a search of up to 100 bp s^{-1} at $37°C$.[160,349,350,351]

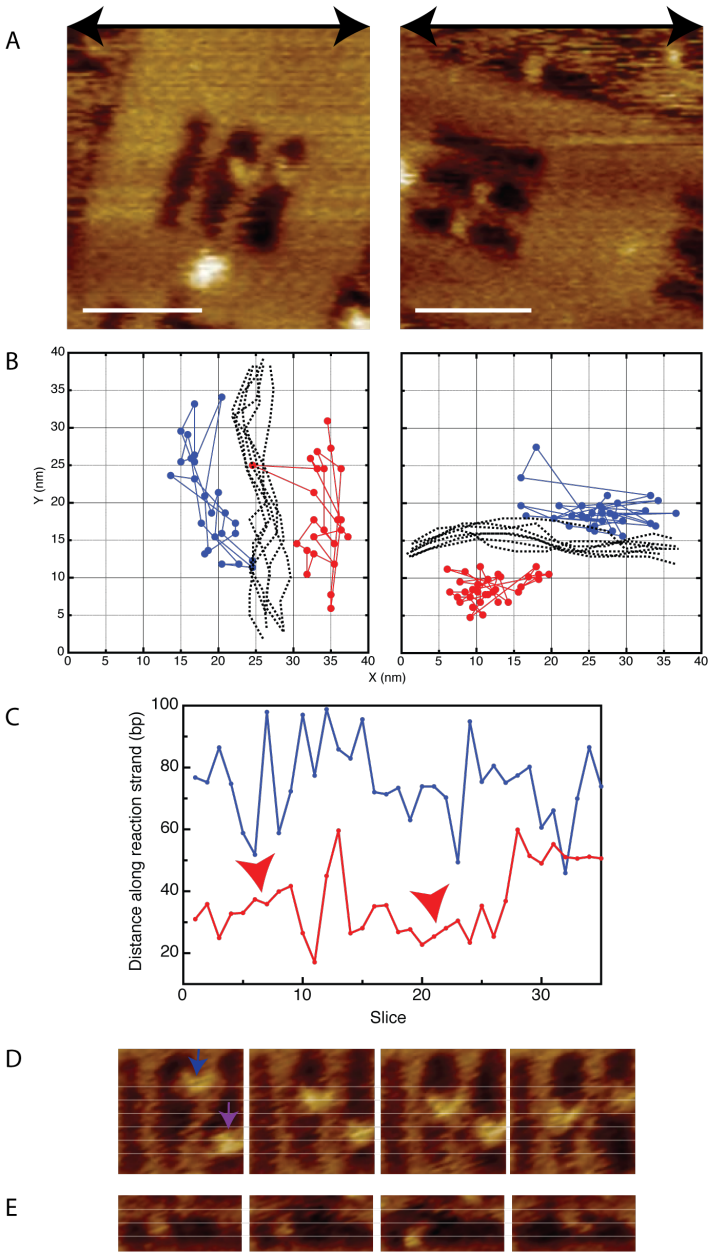

Figure 8.10: ***Observation of NPF homology searching dynamics within DF2S using HS-AFM. A)*** *The dynamics of 30 nt NPFs are observed in several orientations, perpendicular to and parallel to the scan axis (black arrows).* ***B)*** *Tracking the centre of mass of the NPFs with respect to the DNA strand reveals similar motions irrespective of the scan direction.* ***C)*** *A representative plot showing the NPF distance along the DNA from the polarity through successive images. The NPFs are observed to jump back and forth, suggestive of a random sampling mechanism. Short regions demonstrate progressive "slides" along the DNA (red arrows), representative of one dimensional facilitated diffusion. The distinction of these events can clearly be seen in the HS-AFM images demonstrating facilitated diffusion (sliding)* ***(D)*** *and random sampling (hoping)* ***(E)****. Scale bars (A) = 40 nm. Z scale = 6 nm.*

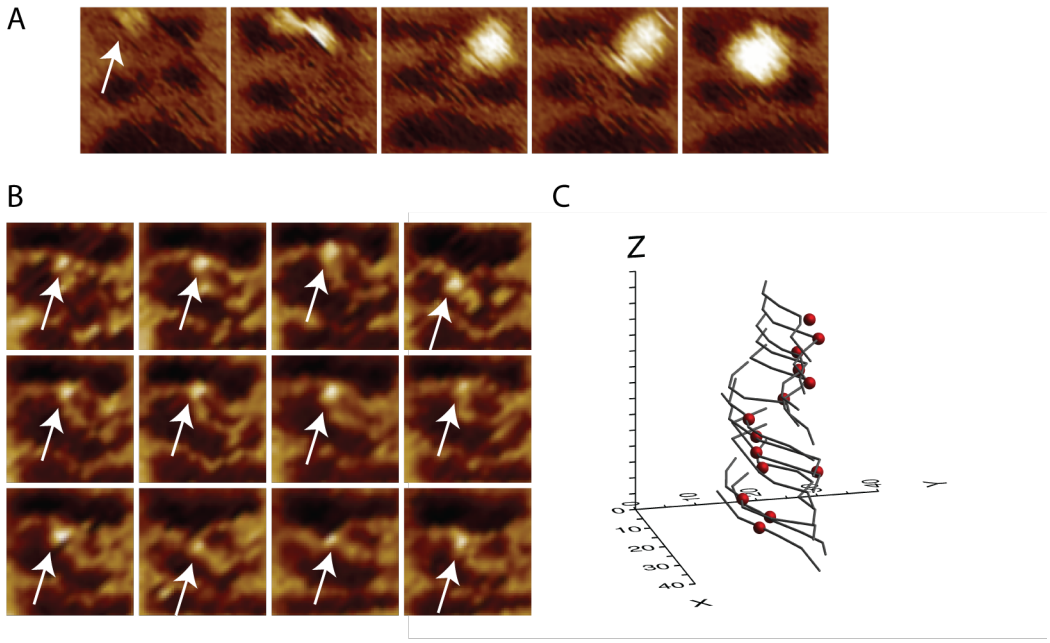

Figure 8.11: ***Observations of correct patterning in dynamic experiments with the HS-AFM.A)*** *A 30 nt NPF is seen to arrive from solution and slide quickly to stall at the region of homology.* ***B)*** *Another NPF (white arrows) is observed to consistently move with the DNA through consecutive images whilst centrally positioned at the region of homology. This is depicted clearly in the vector plot in* ***(C)****. The centre of mass of the NPF (red sphere) in* ***(B)*** *is plotted in relation to the DNA strand (grey lines) over time (Z axis). Image size = 40 nm. Z scales, (A) = 6 nm, (B) = 5 nm.*

From the observations made here, an overall search rate of 3.26 bp s^{-1} $\pm$ 0.53 bp is calculated under the conditions conducted in these experiments. Interestingly, this is made up of a combination of 5.73 bp s^{-1} $\pm$ 0.47 bp for hops and 1.1 bp s^{-1} $\pm$ 0.18 bp for slides. The discrepancy between the proposed and the observed rates, is nearly a 30 fold difference and is largely attributed to the experimental setup, where interactions are conducted at $22°C$, and samples are partially immobilised at a solid–liquid interface. Hence this difference is considered a function of the severely reduced rates of NPF diffusion at the surface – as seen previously in section 7.2.9 – and in solution. Furthermore, the imaging rate of these experiments must be factored in, typically 3 seconds per image in these experiments. Although far in excess of standard AFM frame rates, this is still limiting with respect to observing protein dynamics and is obviously not "real-time", hence one must consider that only a fraction of the interaction information is captured which may lead to slides being misleadingly observed as hops, for example.

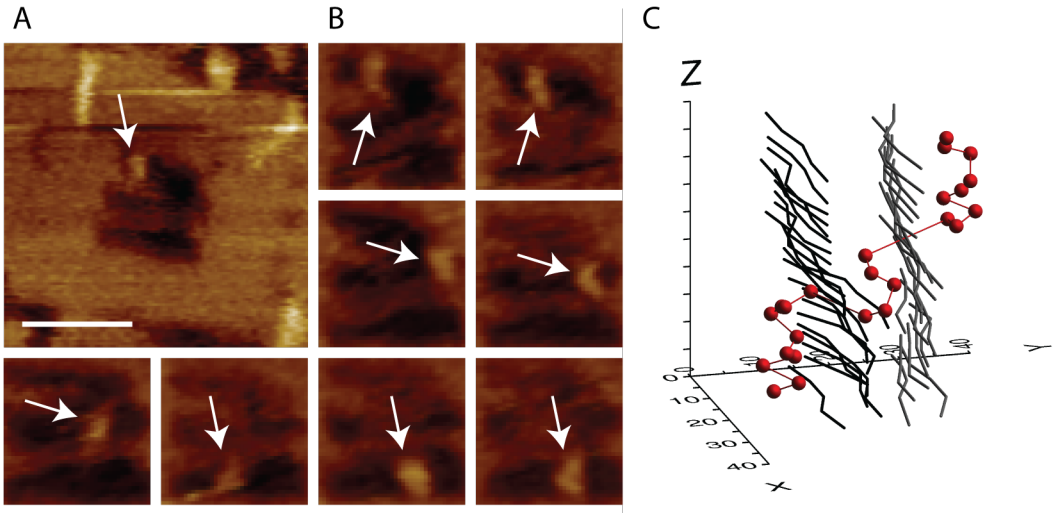

Figure 8.12: ***NPF homology searching over non-contiguous regions of DNA. A)*** *The position of a 30 nt NPF (white arrow) is shown relative to the two central strands within a DF2S frame.* ***B)*** *A series of HS-AFM depicting the NPF sliding along the top DNA strand and transitioning to slide along the bottom DNA strand. These events are shown clearly in the vector plot in* ***(C)****. The NPFs centre of mass is plotted (red sphere) in relation to the DNA strands (grey and black) over time (Z axis). Scale bar = 40 nm. Z scale = 6 nm.*

8.2.10 Observation of Homology Location

Although searching behaviours are regularly observed throughout these experiments, successful location of homology - and hence transition from synaptic to postsynaptic joint - has been difficult to capture. One may consider that despite clever experimental design, the HS-AFM is only able to observe a fraction of the sample substrate at any given time. Hence, catching interactions at the right time remains difficult, particularly given the low concentration of interaction species used here.

Nonetheless, evidence for these can be seen in figure 8.11 A, where a NPF appears directly from solution, slides along the dsDNA and stalls at the region of homology. Furthermore, other observations are suggestive of the existence of post-synaptic complexes where NPFs are consistently located at the same point on the DNA through successive images (figure 8.11 B & movie 8.2). This is demonstrated in figure 8.11 C, where the NPFs centre of mass is tracked relative to the DNA over time (Z axis). From this, the central location and stable nature of these complexes is suggestive of a transition to a postsynaptic complex at the region of sequence homology.

8.2.11 Searching of Non-contiguous DNA Sequences

Interestingly, NPFs were seen to search both the reaction and control dsDNA strands with no discernible differences, suggesting continuous homology probing of all available sequences. Note that although possible, no evidence of NPF interaction with portions of the DNA frame – folded or partially folded – were observed. Moreover in rare cases, NPFs were observed to search non-contiguous regions of DNA, in this instance swapping between the two central strands (figure 8.12 & movie 8.3). This is interesting as it can be considered that such swapping between separate regions may support the involvement of inter-segmental transfer, as suggested by others.[11]

Such behaviour has previously been proposed for RecA NPFs, where by distal portions of a long NPF are able to sample separate regions of the same DNA simultaneously. As such, it has been proposed that inter-segmental transfer may result in faster template searching or the ability to circumnavigate regions of persistent confinement or complex topology.[11] In the case of this work, the NPFs utilised here are considered too short to undergo true inter-segmental transfer, representing only 1.5 turns of the nucleoprotein helix. Nonetheless these observations are interesting with regards to understanding the possible involvement of this mechanism in the RecA mediated homology search.

8.2.12 A Specifically Orientated Interaction

As could be seen from the images in figure 8.10 A, it appears that it is the termini of the NPFs that is interacting with the DNA strand in a specific orientation. Intriguingly, as the geometry of the DF2S frame is precisely known, it can be used to estimate the angle of this interaction, where the NPF is assumed to lay within the surface plane. This is not an erroneous assumption, given the presence of Mg^{2+}species within the imaging environment which are capable of mediating a transient binding between two negative planes (see section 7.1.1) – In this case the mica and the negatively charged C - Terminus of the RecA (see section 3.2.2). The plot in figure 8.13 A depicts the distribution measured angles of orientated NPF interactions. This is calculated by measuring the angle between the long axis of the NPF with respect to the direction it is travelling along the dsDNA (figure 8.13 B).

From this, a characteristic interaction angle of $111^\circ \pm 2.7^\circ$ between the NPF and dsDNA is found. This orientated interaction is highly suggestive of the dsDNA travelling aligned within the NPF helical groove, shown schematically in figure 8.13 C.

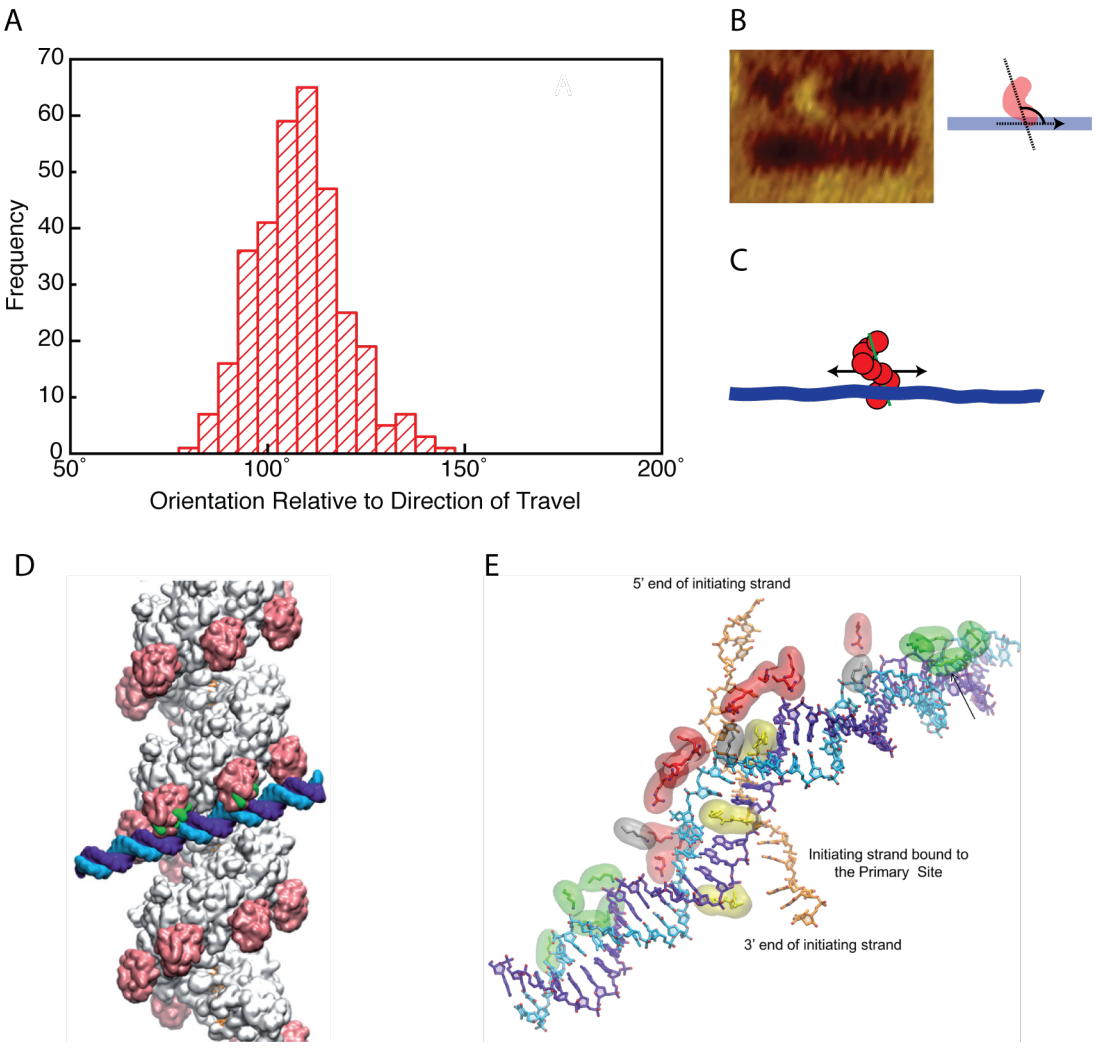

Figure 8.13: ***NPFs are observed to undertake an orientated slide along DNA. A)*** *A histogram plot depicting the angle of interaction of the NPFs as measured in* ***(B)****.* ***C)*** *The observed interaction of the NPFs is shown schematically.* ***D & E)*** *These observations are found to be in-line with the orientated interaction regimes proposed by the molecular dynamic simulations conducted by Yang et al.*[160] ***D)*** *The incoming dsDNA aligns into the helical groove of the NPF through interactions with the C terminal domains (red).* ***E)*** *The encapsulated ssDNA is shown in orange, with the interacting dsDNA shown in purple and blue. The residues responsible for DNA interactions are highlighted for the C terminal domain (green), binding site II (red) and L2 binding loop (yellow). Figures* ***(D)*** *&* ***(E)*** *are adapted from Yang et al.*[160]

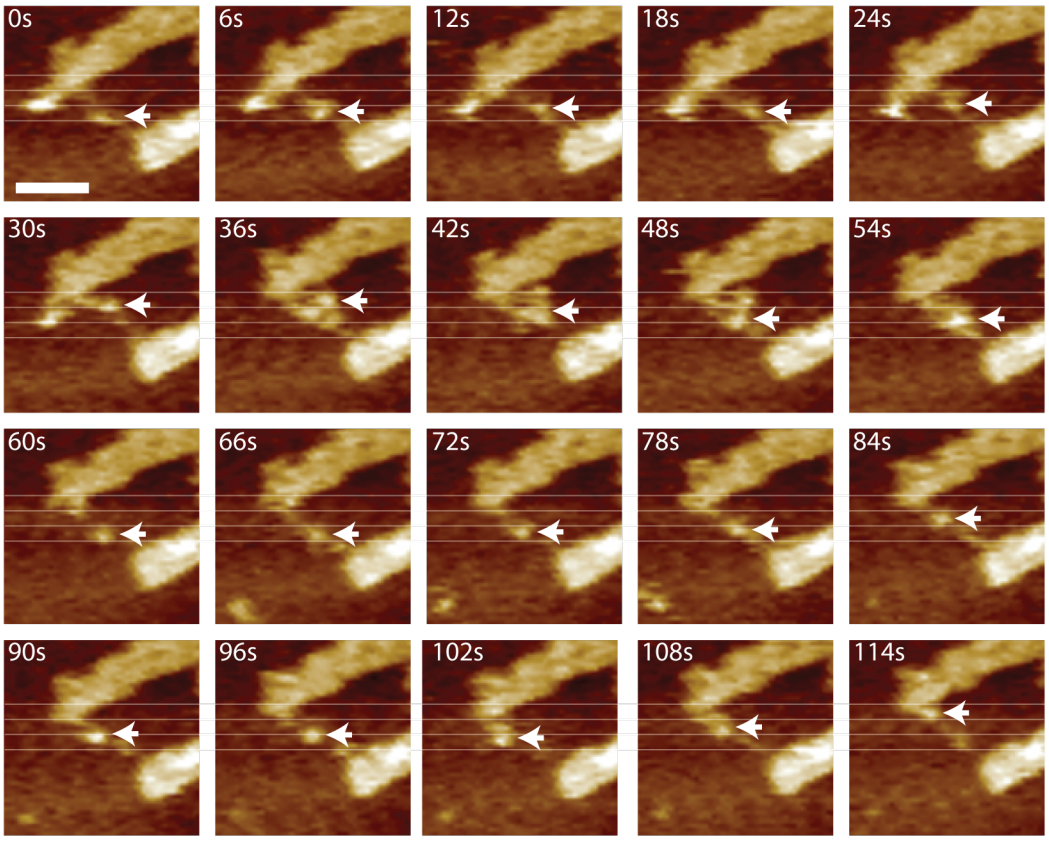

Figure 8.14: ***Observation of long distance NPF slide in broken DF2S frame.*** *A series of sequential HS-AFM images depicting the long distance slide of a 30 nt NPF (white arrows) within a broken DF2S frame. A second NPF is seen to arrive and briefly interact (36 s). Scale bar = 40 nm. Z scale = 6 nm.*

This fits well with the interaction geometries proposed by Yang et al, through molecular dynamics simulations (figure 8.13 D & E), appearing remarkably similar.[160] Here, the dsDNA interacts with the favourably charged C terminal domains (red in D and green residues in E). This is proposed to align the incoming dsDNA enabling it to be transiently bound into binding site II (red residues in E). Note the proposed long axis of the NPF – observed as the encapsulated ssDNA in (E) – with respect to the incoming dsDNA in figure 8.13 D & E. From here, the incoming dsDNA is probed for sequence homology where the residues of the L2 binding loop (yellow in E) intercalate and locally open the DNA (see section 3.2.3).

Although the observations made in this study do not have sufficient spatial resolution to infer the exact geometry of DNA species within this complex, it remains the first direct experimental evidence in support of a transient alignment of dsDNA within the NPF helical groove during homology searching.

8.2.13 Expansion of the Frame - the Development of DF2L

Despite these detailed insights into the RecA homology searching mechanism they are hard to reconcile within the context of a genome wide search, due to the restrictive geometry of the DNA frame - where the 30 nt NPF represents roughly 1/4 of the available sequence space available. As discussed previously (section 6.2.1) it is not practical to reduce the dimensions of the NPF in these experiments any further and hence the only option is to increase the size of the DNA frame.

In further support of this, observations made in partially unfolded or mechanically broken DNA frames - for example in figure 8.14 - reveal longer distance sliding behaviours than suggested in the previously described data. In this example, part of the scaffold is loosely strung between two remaining folded regions of the DNA frame (figure 8.14). Here, a NPF (white arrow) is observed to slide the length of the strand prior to interacting with a second incoming NPF (see figure 8.14 & movie 8.4). With respect to this, it was concluded that an expansion of the DNA frame was required for more biologically relevant homology searching observations to be made. This would enable the characterisation of much longer sliding mechanisms, possible template looping – where the DNA was made flexible enough – and the future ability to conduct a full strand exchange reaction within the DNA frame window.

In order to increase the DNA frame efficiently, a collaboration with Masayuki Endo (Sugiyama Group, Kyoto University) was undertaken in order to draw on the groups' depth of expertise in this area. It is important to note that the frame design was jointly developed between Prof. Endo and

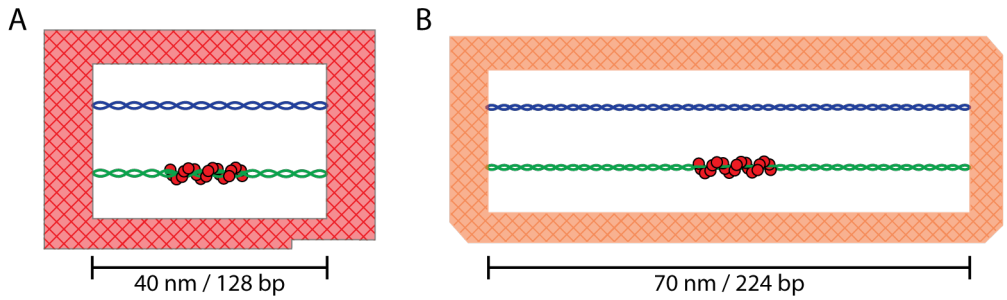

Figure 8.15: ***A schematic diagram comparing the DF2S and DF2L frames.*** *The DF2L frame enables nearly twice the sequence to be hosted within the central cavity for conducting RecA homology searching experiments.*

the author, from a base design taken from previous work in the Sugiyama group. Once the design was finalised, the frame was folded in the Sugiyama group and shipped to the Bioelectronics group, where the central strands were developed and incorporated.

Within the Sugiyama group, a larger frame already existed which provided a suitable base for re-development for this work. At this stage, the frame was modified through several iterations to introduce two parallel sets of anchor points - as utilised in the previous smaller design. It was concluded that the frame could be expanded to harbour 2 DNA strands with a maximum sequence length of 192 bp. This was achieved by recessing the anchor loops into the frame walls (see the schematic in appendix, figures A.2). This larger frame is referred to as DF2L.

However, this solution was quickly found to be problematic, where steric hindrance from the frame walls prevented successful incorporation of the central dsDNA strands in the initial trials. As such, the frame was re-designed to marginally increase the cavity further in order to reveal the binding loops, adding bracing to the internal corners to achieve this. In addition, this increased the length of the central strands to 224 bp. The design iterations of this frame can be seen schematically in the appendix (figures A.3 and A.4) and a scaled comparison of small and large DNA frames is given in figure 8.15, indicating the significant size increase.

8.2.14 Generation of the DF2L Central Strands

During the initial development of DF2L, the central strands were designed. These sequences largely followed the format of those used in DF2S, including two different length versions, at 224 bp and 234 bp. Due to their size, it was not possible to simply anneal these from two oligonucleotides and hence the sequences were synthesised into a plasmid, allowing their amplification by PCR (see methods 5.2.1). However, due to subsequent changes in the frame

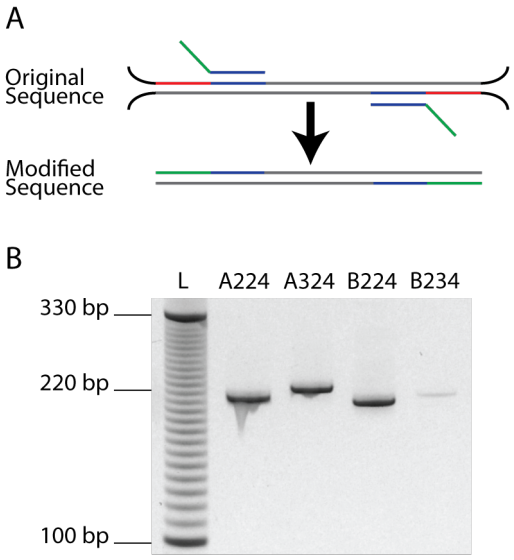

Figure 8.16: ***The introduction of new anchor sequences to the DF2L central sequences. A)*** *A schematic diagram depicting the trailing PCR amplification method. The old anchors (red) are replaced by new anchors (green) which are incorporated as 5' trailing sequences, with the 3' region (blue) priming the PCR reaction on the original sequence* ***B)*** *A 10% polyacrylamide gel depicting the successful amplification of 3 of the 4 DF2L central strands using trailing PCR. Poor yield is observed for B234 for an unknown reason.*

design, a trailing PCR method was applied to introduce the new anchor sequences during amplification (this is detailed in methods 5.2.2).

This approach was successfully applied to re-generate 3 of the 4 central dsDNA sequences with high yields, referred to as A224, A234 and B224 (figure 8.16). The designation A or B refers to the pair of anchor points in the DF2L frame to which they bind, top and bottom (with respect to the polarity marker), respectively. The reason for the apparent failure of B234 is unknown. However, as the 224 bp sequences were to be utilised in the first instance and the 224 and 234 bp sequences only differ by 10 bp added evenly before the anchors, no efforts were made to rectify the PCR reaction at this point.

Prior to further processing of the central strands, both sets of central sequences – A and B – were subjected to a restriction assay to check the patterning efficiency, as described previously for DF2S (section 8.2.4). A224 and B224 were patterned using respective 30 nt NPFs, where successful patterning at a 3:1 (NPF:DNA) ratio (figure 8.17) was concluded to be sufficient for use in the DNA frame. Note that the sequence A224 and A234, differ only in 10 bp added evenly to either termini – following the anchor sequences – and hence the restriction site and region of homology remain identical. As such, the patterning efficiency of A224 was assumed to be indicative of A234 also.

Following the successful generation and characterisation of dsDNA, 16 nt ssDNA overhangs were generated in order to incorporate the central strands into the origami frame. These ssDNA regions were generated at the termini by nicking the template dsDNA with the enzyme nt.bst.NBI and competing away the redundant ssDNA. For details of this process please refer to the methods

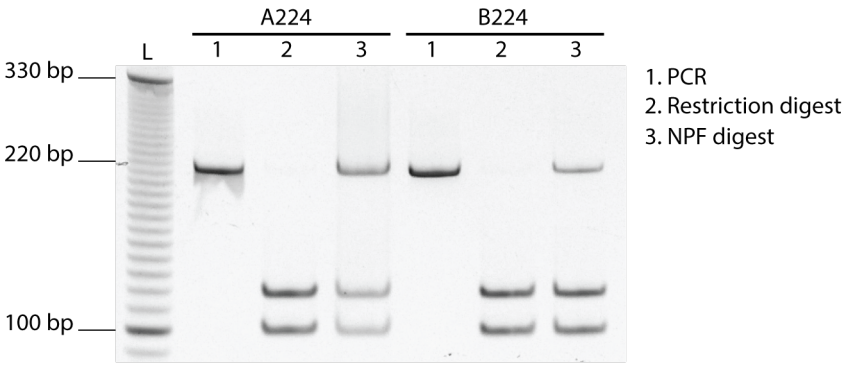

Figure 8.17: ***Restriction assay assessment of DF2L central strand patterning.*** *A 10% polyacrylamide gel demonstrating successful patterning and protection of the central strand sequences A224 and B224 templates by respective homologous 30 nt NPFs.*

(methods 5.2.5). This was found to be 100% successful for A234 and B224, but interestingly, only 50% efficient for A224. Again as with the trailing PCR reaction, discrepancies between the reaction yields was evident – considering here that the nicking site and anchor sequences for A224 and A234 are identical, this difference is curious – but was nonetheless tolerable for the subsequent incorporation into the DNA frame.

8.2.15 Incorporation of DF2L Central Strands

Subsequent incorporation of the central strands was undertaken, as described for DF2S. Here, all three central strand variants were incorporated individually following the procedure discussed previously (see section 8.2.5), conducted here at a 5:1 (central:frame) ratio with an average yield of 60%. Despite the re-designed DF2L containing exposed binding loops, a consistently higher incorporation yield was not achieved despite increasing the ratio to 10:1.

Figure 8.18 shows representative AFM images of DF2L frames with singularly incorporated central strands, A224 (A & B) and A234 (C). It is important to note the large degree of flexibility that these central strands are afforded within these frames, particularly when considered with respect to those observed in DF2S. This is most prominently observed for the incorporated sequence A234 (figure 8.18 C), where the central strand is found to be difficult to track. This became a prominent issue when NPF interactions were observed in this DNA frame.

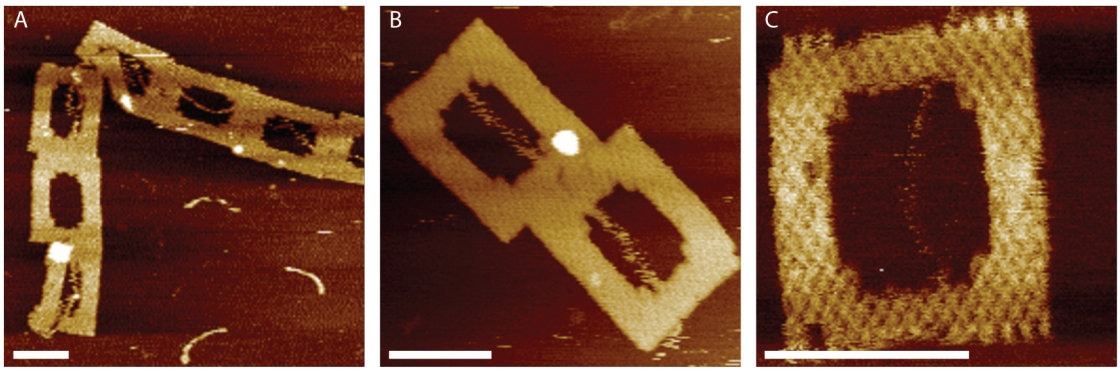

Figure 8.18: ***A montage of AFM images depicting the preparation of DF2L and the incorporated strands.*** *The frames are observed to be well formed and the flexibility of the incorporate strands is evidenced by the difficulty in tracking them. Scale bar = 70 nm. Z scale = 3 nm.*

8.2.16 Nucleoprotein Filament Interactions within DF2L

NPF observations in DF2L were conducted using the conditions established previously for DF2S. Here, 30 nt NPFs were introduced to DF2L frames following their deposition in Tris-Mg buffer (methods 5.1.3) containing 10 mM Mg^{2+}(methods 5.2.8).

However, despite the previous successes with DF2S, only a handful of successful observations were made with DF2L in the time available. This is largely attributed to the high degree of flexibility of the frame itself – due to the reduced frame structure. DF2L frames were seen to flex and "breath" between successive images in a series, which in turn causes the distance between the binding loops to shrink making the central dsDNA more flexible. As a consequence it became particularly difficult to observe the central strands within these DNA frames (figure 8.19 A & movie 8.5). Moreover, this reduction in structural integrity manifests itself more profoundly as DF2L frames are found to be more susceptible to damage by the HS-AFM probe and hence resulted in large losses during observations.

However, despite these setbacks, some observations of NPF searching behaviours were made. Figure 8.19 B shows a representative series of HS-AFM images depicting the progressive movement of a NPF along the expected trajectory of the dsDNA. Although, as discussed, it is difficult to observe the dsDNA in these images, there are particular regions at the right-hand side of the reaction window (figure 8.19 B, top right of the image) which are highly suggestive of the presence of the dsDNA. Therefore it is suggested that these observed NPFs are indeed conducting a homology search, although it is difficult to say whether the observed motion is that of a sliding or a hopping regime. This apparent difficulty in tracking the DNA substrate may be considered

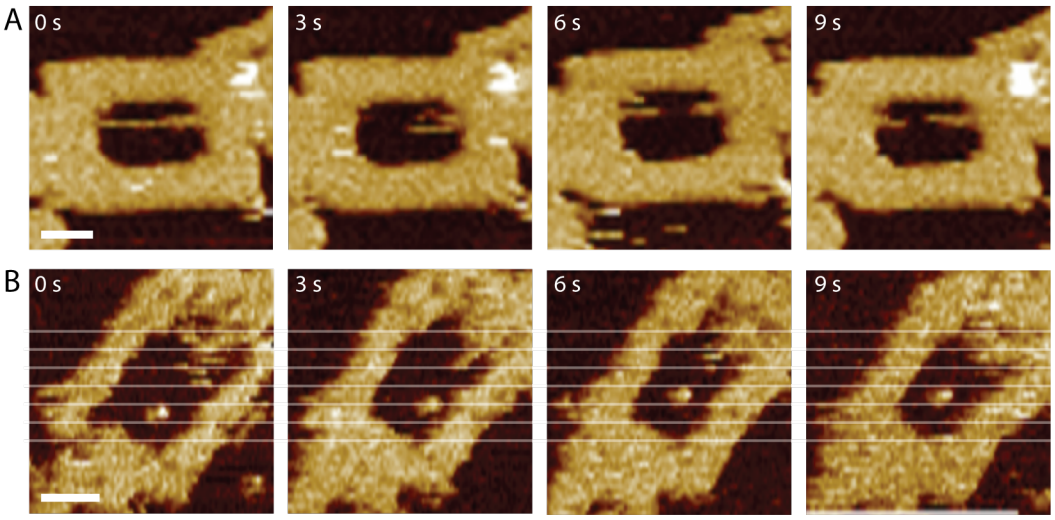

Figure 8.19: ***A series of sequential HS-AFM images depicting the initial observations of NPF homology search conducted within the DF2L frame. A)*** *The frames and central strand are found to be highly flexible resulting in difficulty resolving the DNA.* ***B)*** *An NPF is observed to slowly and progressively migrate along the long axis of the DNA frame, suggestive of sliding on the DNA. The presence of the central strand is suggested in the top right of the image. Scale bar = 35 nm. Z scale = 6 nm.*

a step in the wrong direction, resulting in similar quantification problems as those experienced in the previous chapter.

8.3 Conclusions

The work in the previous chapters had successfully demonstrated the observation of RecA NPF interactions with dsDNA. However, limitations in the experimental approaches had limited the meaningful quantification of the intrinsic orientational or positional information.

In order to circumnavigate these experimental limitations, the work in this chapter adopts the DNA origami frame methodology of Sugiyama and Endo et al[128,236] to further explore the RecA mediated homology searching mechanism. DNA origami frames are able to act as a positional and orientational reference as well as to anchor incorporated dsDNA templates to a mica substrate. Here, a DNA frame design was adapted from the literature and successfully utilised for these purposes.

Observations of NPF interactions with dsDNA conducted within DF2S have provided further support for the existence of facilitated diffusion. Here, NPFs are clearly observed to undertake slides of an average of 10 bp along the dsDNA template with a measured rate of 1.1 bp s^{-1} $\pm$

0.18 bp. This is accompanied by distinct hopping or random sampling – at 5.73 bp s^{-1} $\pm$ 0.47 bp – which appears to dominate the NPF interaction behaviour. Note that these are conducted at $22°C$ with the species partially immobilised upon a surface and are unlikely to be representative of physiological rates.

Despite the clear observation of transient homology probing interactions, limited evidence for the transition of a synaptic to a postsynaptic complex is evidenced. Nonetheless, successful homology searching interactions have been demonstrated, typically seen in DNA frames which were not continuously observed. This highlights that despite excellent experimental setup, one must observe in the right place at the right time. Further, it must still be considered that the presence of the surface is likely to play a large role in limiting the full wrapping of a target DNA molecule into the NPF complex.

Moreover, detailed information regarding the orientation of the NPF interaction has been evidenced for the first time, in support molecular dynamics simulations conducted by others.[160] Here, NPFs are revealed to align with a typical angle of 110° – not previously reported – enabling the dsDNA to reside within the negatively charged tract of the helical groove, which is highly suggestive of transient associations of the dsDNA occurring within.[160]

Despite these successes, the limited geometry of the DF2S was deemed restricting and attempts to move to a larger DNA frame in collaboration with Masayuki Endo (Sugiyama group, Kyoto University) were detailed. Such a larger reaction space would allow for further experiments to be orchestrated in the future, such as the full RecA mediated strand exchange. However, although the DF2L frame was successfully developed and initial observations conducted, the fragile nature of the frame was found to be a problem and hence limited the number of successful observations carried out within the time constraints of the project.

With regards to future development, it is proposed that the DF2L frame should be further developed utilising a longer ssDNA scaffold - enabling a more robust structure to be formed. Although uncommon, it has been demonstrated that it is possible to utilise longer ssDNA templates in DNA origami, making the approach inherently scalable provided the availability of template DNA.[72] For example, as M13 is a filamentous bacteriophage, it is possible to increase the length of its DNA drastically without hindering its replication. Furthermore, 8.6 kbp M13 bacteriophage are already commercially available (New England Biolabs, MA, USA) for the development of phage display libraries in molecular biology. Hence, the expansion of DF2L to use a 8.6 kbp template is entirely possible and would enable a more robust structure to be developed.

Thus far in this thesis, the work has attempted to use the AFM to interrogate the dynamics of the RecA mediated homology search. In contrast the next chapter will apply the AFM to push the resolution boundaries on RecA NPFs and its manipulation of DNA, in order to provide further insight - connecting form to function.

Chapter 9

Observing the Manipulation of DNA Structure by RecA

9.1 Introduction

The elucidation of discrete structure forms a central pillar of the study of biological macromolecules, but to date typically relies on ensemble-averaging of crystalline samples (see section 3.4).[19,18,149,22] Such methodologies offer poor sensitivity to the subtle intramolecular variations which are arguably of most importance in connecting form to function. Within this context the AFM commands a unique position, offering the ability to image isolated biomolecules with sub-molecular resolution, present within physiological environments - albeit absorbed at an interface (section 4.1). Thus AFM has been shown to offer further structural insight complementing NMR, X ray crystallography and electron microscopy methodologies.[352,353]

In modern AFM, the position and vertical force exerted by the probe can be measured and controlled with sub-Angstrom and pico-Newton accuracies, respectively.[354] Nevertheless, the probe remains in physical contact with the sample to acquire an image, inevitably resulting in distortion of the substrate. This is of particular concern when imaging soft matter, such as biological molecules and derived synthetic structures – many of which natively respond to external mechanical forces.

The compliant nature of biological samples has previously been addressed through the packing of molecules into 2D arrays, thus increasing their effective Young's moduli.[354,355] This approach was used by Mou et al to demonstrate the helical pitch of DNA in 2D crystals.[356] However, only limited numbers of biological molecules of interest can be assembled into such arrays. Therefore,

in order to resolve sub-molecular structures in topographically heterogeneous biomolecules or biologically derived synthetic structures, the sensitivity and accuracy of the cantilever loading force must instead be addressed, as the spatial resolution is directly related to the applied force (see section 4.3).[357]

One approach to this, demonstrated by Santos et al, was to apply sub-nanometre amplitude modulation using very small set-points (SASS imaging) enabling continuous probe oscillation within an ambient water layer.[358] With this method, they successfully resolved the B-form helical pitch in isolated DNA molecules. In addition, further advances in miniaturisation of AFM cantilevers have drastically improved the signal to noise ratio, which continues to enhance the force sensitivity of the tip-sample interactions.[231,357,359] Where this has been coupled with developments in frequency modulation (FM) – tracking subtle shifts in the resonance of the cantilever due to the tip–sample interactions – routine atomic resolution on hard samples are now achieved.[360] Although FM modes are increasingly available on commercial systems, they have not yet seen wide adoption.

An alternative approach to minimise the impact on the sample by reducing the loading force is to exploit the AFM's ability to acquire force–distance curves, previously utilised for mechanical mapping. Recent hardware advances, in particular the bandwidth of the controller electronics, has led to the development of rapid force-curve-based imaging methods – such as Peakforce tapping – enabling a continuous estimation of the peak loading force applied to the substrate, which is referenced as a baseline for actuating system feedback (see section 4.4.2).[282,285]

This methodology has been shown to allow routine observation of the helical pitch on isolated DNA molecules, conferring similar resolutions to those of advanced AM or FM modes.[361] Here, the loading force can be continually estimated and adjusted in response to sub-molecular features enabling the resolution of local structural variation.[361] In this chapter this rapid force–curve–based imaging is applied to directly interrogate the structure of dsDNA in complex topologies and under the manipulation of RecA.

It is proposed from crystal structure observations that upon binding, the RecA NPF partially relaxes the dsDNA helical pitch from 3.4 nm to 5.2 nm allowing the presentation of the internal base pairing to the encapsulated ssDNA. All things being equal, it is proposed that the relaxation of a dsDNA molecule internal to the NPF complex must be accompanied by a helical compression proximally. Although this compression may dissipate through the DNA within a solution phase, it must be considered that where the rotational freedom is restricted - such as on a surface - that any

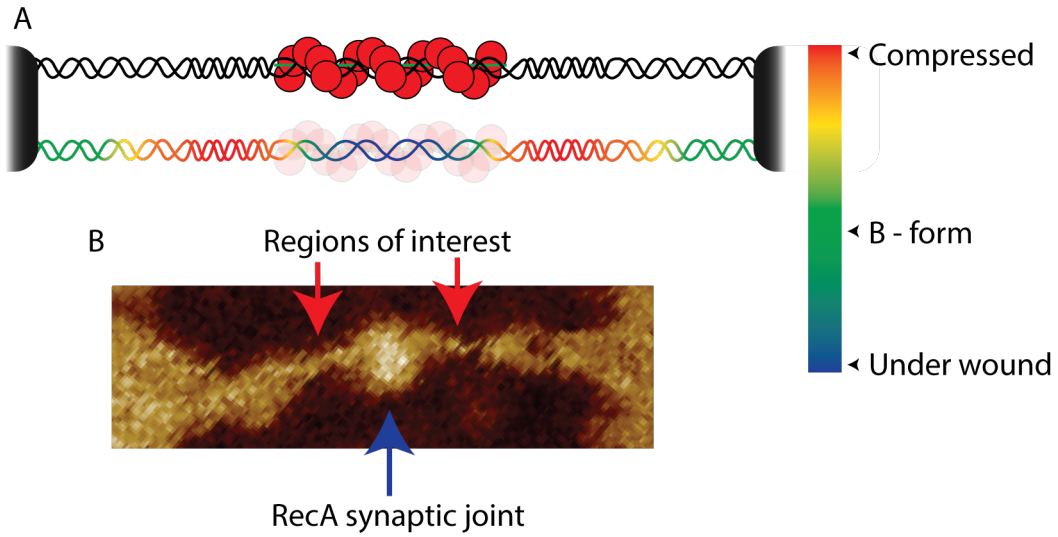

Figure 9.1: ***An overview of the approach taken to trap the RecA induced helical compression of DNA. A)*** *A schematic diagram depicting the trapping of the induced helical compression in DNA. The termini of the DNA molecule have their rotational degrees of freedom restricted through the use of double anchors.* ***B)*** *The regions of interest for imaging the helical pitch of DNA are highlighted on the AFM image (red arrows).*

induced variation in the helical pitch will remain present and might be detectable by AFM.

Building on the work carried out in the previous chapter (section 8.2.1), the DNA origami frame - previously used for dynamic studies - was employed to trap any induced compression of the DNA by substrate RecA (figure 9.1) by securely anchoring the termini. Prior to this, this chapter explores high resolution observations of DNA - both linear and woven within DNA origami - and the RecA NPF itself, in isolation.

9.2 Experimental

9.2.1 Resolving the Helical Pitch of DNA

Looking towards imaging the helical pitch of DNA in complex topologies and under the manipulation of RecA, reliable imaging of the B - form helical pitch must be established on native DNA in isolation.

For these experiments, dsDNA was bound very tightly to the mica support surface to minimise movement of the molecule during probe interactions. This was in stark contrast with the work described in the previous chapters. Here, DNA was immobilised in aqueous buffer containing 10 mM $Ni^{2+}$and incubated on the surface for up to 1 hour (see methods 5.2.8). Further incubation

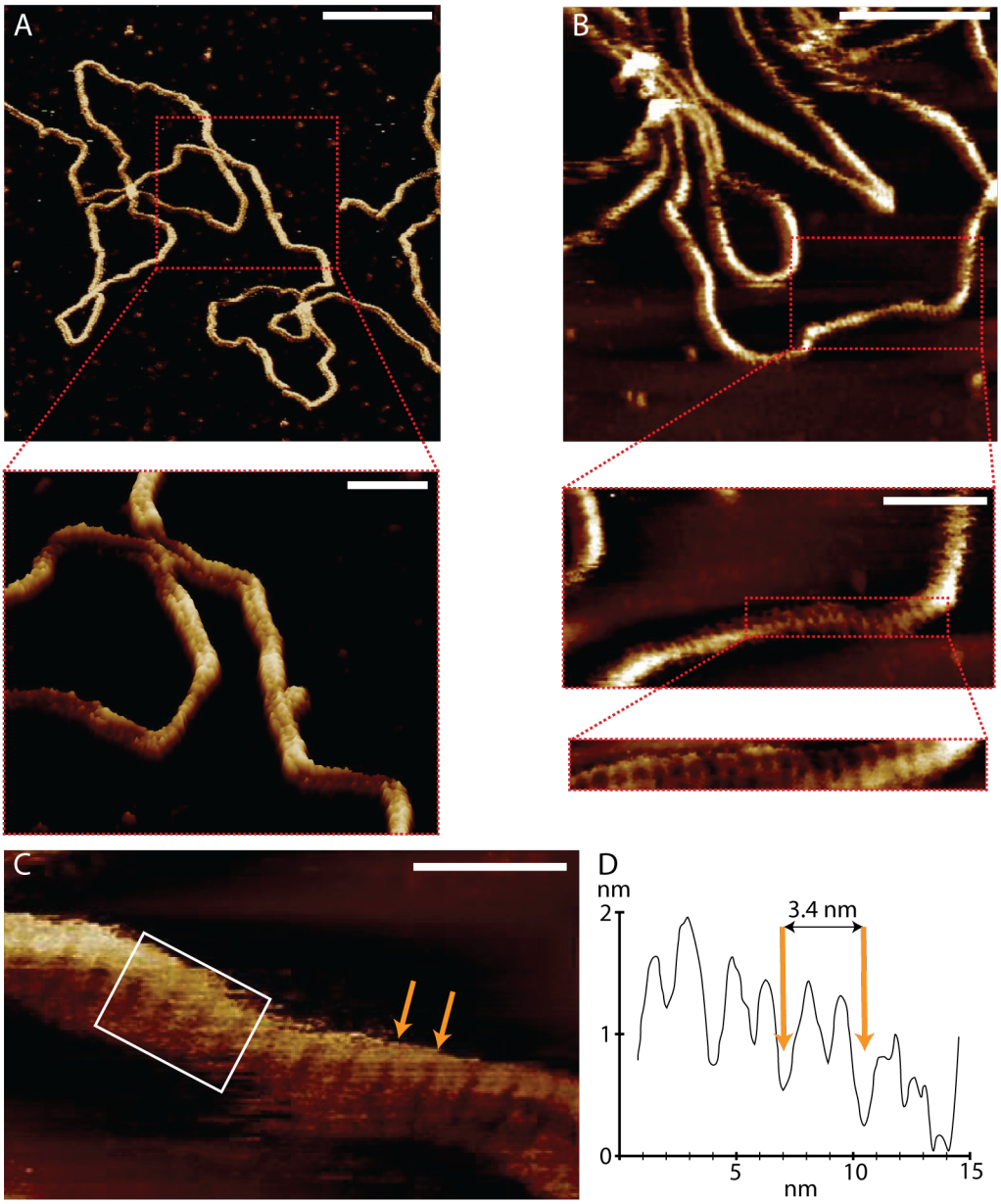

Figure 9.2: ***A montage of AFM images depicting the helical pitch of DNA observed with rapid force-curve-based imaging.*** ***A)*** *The application of 100 pN of force minimises the compression of the DNA molecule. The helical pitch is observed along the length of the molecule along with long scale periodicity, representing supercoiling.* ***B)*** *The helical pitch is most clearly resolved where the DNA is tightly bound to the surface. Note the loosely resolved regions.* ***C)*** *Both the major and minor grooves are clearly resolved, demonstrating the characteristic 3.4 nm helical pitch of B form dsDNA. Scale bars, (**A & B**), top = 50 nm, bottom = 15 nm. (**C**) = 15 nm. Z scales = 3 nm.*

or higher cation concentrations tended to lead to abhorrent surface crystallisation and indeed, this effect limits the imaging duration per sample. Note that following incubation, the buffer volume was increased to 200 μl diluting the Ni^{2+}concentration to approximately 5 mM throughout imaging (methods 5.2.8).

Samples were imaged using a rapid-force-curve imaging approach (see methods 5.2.8), utilised in-line with recent demonstrations on isolated DNA plasmids by Pyne et al.[361] From figure 9.2 the structural features of the dsDNA samples are clearly resolved where a consistent tip loading force is applied – estimated at 100 - 150 pN here. Under these conditions DNA molecules were routinely observed with heights of 1.9 nm $\pm$ 0.2 nm, representing negligible disruption to the structure by the interaction of the AFM probe.

Furthermore, throughout the structure large scale height variations are observed (figure 9.2 A, and inset), indicating the occurrence of supercoiling within the λ DNA template. This information is interesting as it highlights the possibility to directly study strain throughout the DNA molecule, which may be of crucial importance to characterising complex DNA structures.

Upon closer examination of the DNA, a distinct banding becomes immediately apparent throughout the molecule (figure 9.2 B, and inset). Importantly, this banding appears in multiple orientations, running perpendicular to the long axis of the DNA. Figure 9.2 C shows a magnified region demonstrating the distinct deeply grooved banding (orange arrows). Cross sections taken along the DNA reveal the B–form helical pitch of 3.4 nm, with both the major and minor grooves clearly resolved (figure 9.2 D). Interestingly, under an estimated applied load of 100 pN the typical groove depths are measured to be 0.8 nm $\pm$ 0.1 nm and 0.5 $\pm$ 0.15 nm for the major and minor grooves, respectively. These values change significantly where the estimated load is reduced below 80 pN indicating a loss of interaction with the fine structure of the molecule (data not shown) – consistent with the findings of Pyne et al. Moreover, from the cross section plot in figure 9.2 D, the large scale periodicity of the molecules super-coiling can be noted in the observed slope.

Given the desire to observe the helical pitch of DNA within constrained and complex geometries, the focus turned to interrogating the complex weave of DNA origami. Attaining these resolutions in the constrained geometry of this complex DNA topology was reasoned as a stepping stone toward imaging the DNA ultra structure juxtaposed at the junction with the RecA nucleoprotein complex.

9.2.2 The Helical Pitch of DNA Observed in Complex Topologies

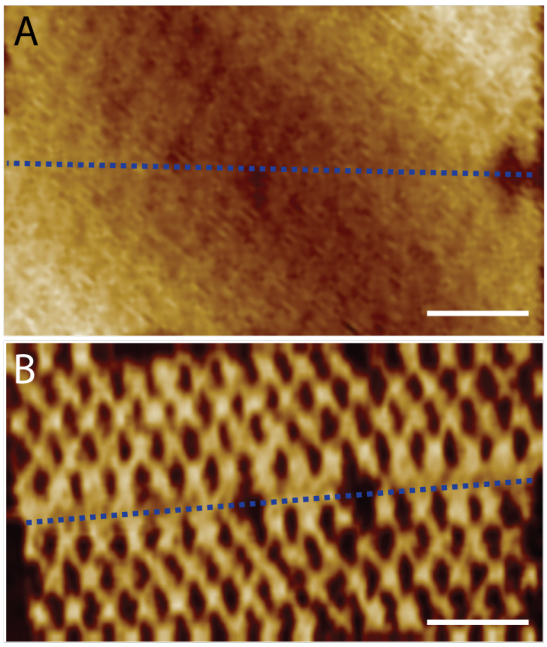

Figure 9.3: ***DNA origami observed with tapping and rapid force-curve-based imaging modes.*** *Typical AFM images taken of the junction (blue line) between two DNA origami tiles.* ***A)*** *The herringbone weave resolved using amplitude modulated intermittent contact (tapping) mode.* ***B)*** *The considerable resolution enhancement achieved through the controlled application of force using a rapid force-curve-based imaging mode. Scale bars = 20 nm . Z scales = 3 nm.*

To date, due to its relatively simple operation and wide spread adoption, tapping mode (specifically, amplitude modulation) AFM has typically been used to characterise low dimensional DNA origami structures.[2] However, as discussed (section 9.1), the resolution of these observations are usually limited by the relatively large loading forces employed during routine operation. From figure 9.3 A, a DNA origami tile is resolved as a solid object when imaged as described. This limited resolution is a consequence of the flattening – and hence broadening – of the individual constituent DNA strands under the typical applied tip-sample load, emphasising the importance of taking into account the mechanical properties of the sample in order to resolve sub-molecular structures.[354]

Figure 9.3 shows several atomic force micrographs of the same DNA origami imaged with amplitude modulation (A) and rapid-force-curve imaging (B). A pronounced difference in spatial resolution is evident, with figure 9.3 B clearly revealing the porous weave of DNA origami across the interface between two neighbouring tiles (blue lines). The individual constituent DNA strands are clearly observed with high definition, and are seen forming the characteristic herringbone pattern, which is merely hinted at in figure 9.3 A. This is achieved by continually maintaining an estimated peak loading force of 100 pN throughout the image acquisition – in-line with the previous investigations on linear dsDNA (see section 9.2.2) – enabling minimal compression of the constituent DNA molecules.[361]

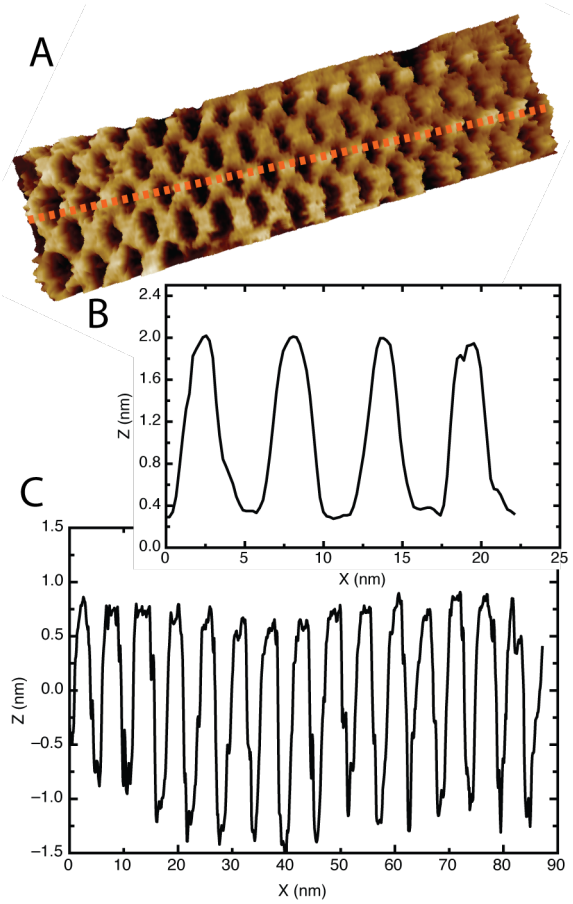

Figure 9.4: ***Rapid force-curve-based imaging enables a true depiction of woven DNA structure*** ***A)*** *An AFM image of a DNA origami tile depicting the position of cross sections shown in* ***(B)*** *and* ***(C)****.* ***B)*** *The individual DNA strands are observed with an average height of 2 nm, demonstrating non compliant imaging of the origami sample.* ***C)*** *Longer cross sections reveal higher order warping in the origami structure as a result of the placement of non-integer cross-overs. Z scale = 3 nm.*

Images with similarly high resolution are obtained when investigating the central regions of the origami tiles (figure 9.4), and as in figure 9.3, the structure of the individual DNA molecules is preserved. From the cross sectional height profiles extracted from the image (taken along the sample as indicated by the red line), it can be seen that the average height of the DNA strands is found to be 1.8 nm $\pm$ 0.3 nm (figure 9.4 B), indicating only minimal compression from the probe interaction. The measured width of the individual DNA molecules in the origami was 2.9 nm $\pm$ 0.4 nm, which is slightly broader than the expected 2 nm radius of B-form DNA. The width broadens further in the vicinity of the vacancies within the weave, suggesting a small amount of probe convolution. Nonetheless, full access to the vacancies is apparent.

Long-distance variations of the height profile of the DNA origami are observed as shown in figure 9.4 C, found to be similar in nature to that seen in linear DNA (figure 9.2). Although these height variations could be a result of subtle differences in the underlying substrate, this is unlikely provided that an atomically flat <001> plane of mica was used as the underlying substrate. Therefore, we suggest that the observed warp is more likely a result of the non-integer helical pitch of B-form DNA (10.5 bp per helical turn – see section 2.1.2) which leads to the accumulation of additional twist throughout successive cross-overs between two adjacent helices.

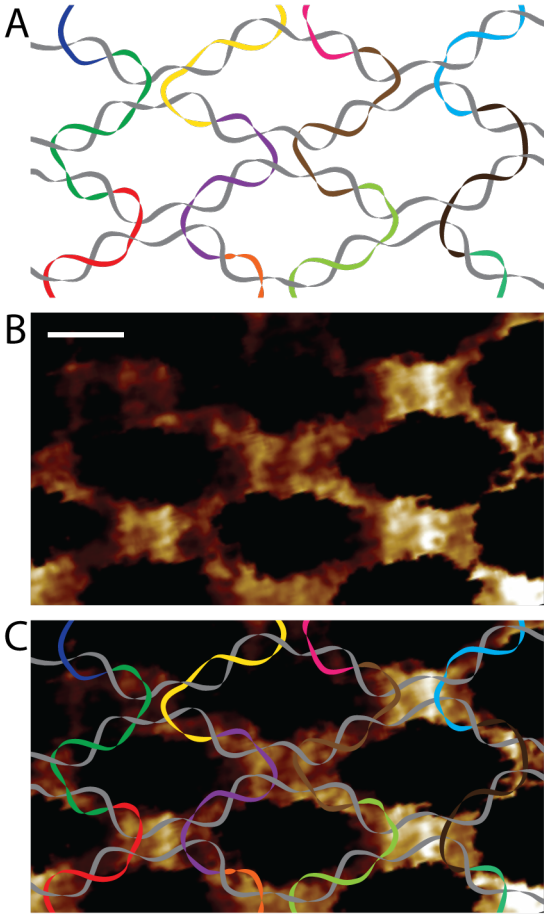

Figure 9.5: ***Confirmation of origami structural observations*** ***A)*** *A schematic diagram depicting weave of DNA strands within an origami structure.* ***B)*** *High resolution AFM image of DNA origami.* ***C)*** *The overlay of* ***(A)*** *and* ***(B)*** *demonstrates the perfect structural replication. Scale bar = 4 nm. Z scale = 3 nm.*

The cross-overs are typically spaced 32 bp, i.e. marginally over three full turns apart. When flat sheets with minimal twist are desired, this over-twist is typically corrected through the inclusion of base mismatches in the staple DNA allowing relaxation (see section 3.1.3). It is important to note that in the DNA origami shown in figure 9.4, mismatches were not included enabling the observation of tile warping. From the images collected in this study, an average tile warp of ± 30 pm out of the imaging plane is observed. However, this is likely to be a significant under-estimate as a result of the electrostatic binding of the origami tile to an atomically flat surface largely neutralising the inherent strain.

Interestingly, the positioning of the crossovers within the origami structure is clearly evident from figure 9.3 B as subtle height variations running in horizontal rows throughout the periodic structure. These crossover points tie neighbouring parallel helices together along the length of the duplex, alternating between either neighbour successively along the helical direction (section 3.1.3). This can be seen in figure 9.5 A, where the grey scaffold strand is seen to run horizontally throughout the structure, with multiple short strands (various colours) running vertically between

neighbouring sections of the scaffold hence, forming crossovers.

DNA molecules are negatively charged, resulting in the repulsion of nearby DNA molecules, and hence the packaging of DNA helices into a crystal-like structure is energetically unfavourable. This repulsion leads to a spatial separation of the helices from one another between crossovers, giving rise to an open porous weave, commonly likened to a herringbone pattern (figure 9.5 A).

This herringbone pattern structure is clearly seen across all the figures presented in this section, and particularly evident in 9.5 B and 9.6 A, which show high-resolution AFM images of selected small areas within the DNA origami structure. For this study, particular care was taken during the immobilisation of the DNA origami onto the mica surface as well as during AFM imaging to ensure that the origami structure is only minimally perturbed, such that only minimal variations from the idealised structure of the origami weave are observed. Figure 9.5 C shows the structural schematic overlaid with the AFM image where it can be seen that they almost perfectly match, confirming minimal perturbation.

The small residual deviations are likely a result of the induced strain as a consequence of the surface-immobilisation of the origami and thus the neutralisation of the inherent warp of the structure. Additionally, small deviations may be the result of residual manipulation by the impact of the scanning AFM probe, which, despite the minimisation of interaction forces, cannot be completely eliminated. Overall, the correlation of the AFM-resolved structure with the theoretically expected structure is remarkable.

Moreover, a distinct banding is evident in figure 9.5 B. These vertical striations are clearly resolved throughout the entire DNA origami structure, running perpendicular to the long axis of the parallel DNA duplexes. Figure 9.6 A shows a similar high-resolution AFM image of part of a DNA origami structure, and the same striations as in figure 9.5 B can be seen. Similar striations were observed on linear DNA with clear alternating banding along the length of the strand (figure 9.2), indicative of the characteristic 3.4 nm helical pitch of the B-form duplex.

This characteristic 3.4 nm helical pitch can be observed for the first time across short continuous sections throughout the woven structure of the DNA origami, despite the more complex topology. Figure 9.6 B shows the height profile taken along the red line of the DNA molecule within the origami as indicated in panel (A). Features clearly resembling the ones observed for linear DNA can be seen, with pronounced major troughs separated by 3.4 nm, indicating the position of the major groove of the DNA double helix. Furthermore, minor troughs located between the positions of the major grooves indicate the resolution of the minor grooves.

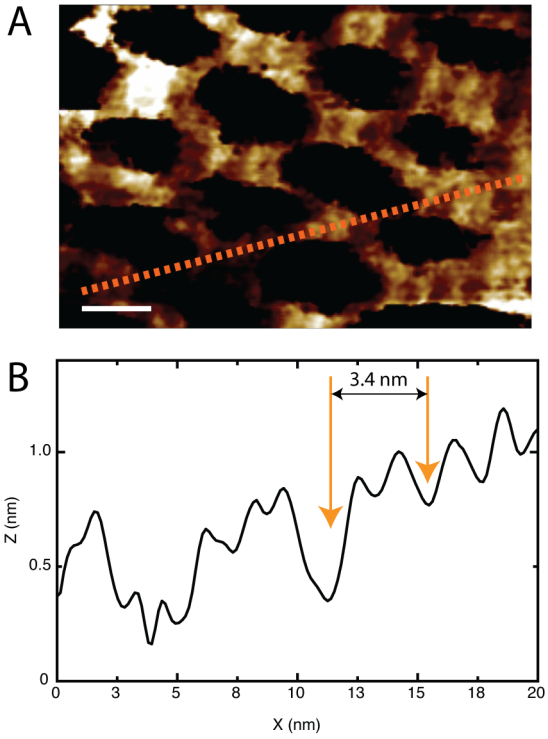

Figure 9.6: ***The helical pitch of DNA observed within complex DNA structures. A)*** *A high resolution AFM image depicting the internal weave of DNA origami. The helical pitch of the integral DNA is observed as banding throughout the structure.* ***B)*** *Profiles taken along the DNA strands (red line) reveal that both major and minor grooves are resolved within this complex topography. Scale bar = 4 nm. Z scale = 3 nm.*

The ability to obtain such high levels of structural details on these complex DNA structures, or even selected regions within them, has significant implications for future designs and applications of DNA origami. It may provide new insight into the inherent strains within the design, porosity of the weave, and longer-scale warp or periodicity within the structure. Such detailed structural and mechanical information is important when developing novel devices from any material, including DNA, and is hence likely to aid the development of DNA structures for specific applications, such as containers for drug delivery systems or where the inclusion of DNA binding proteins is intended.

With respect to the inclusion of RecA, the inherent strain induced by such complex topologies may have detrimental effects upon the efficiency of forming postsynaptic complexes. Having assessed the intricacies of DNA structures - both native and artificial - the next step towards understanding the effects of RecA interaction is to attempt to apply the same resolutions to the RecA nucleoprotein complex itself. This is examined in the following section.

9.2.3 Observing the RecA Nucleoprotein Filament Ultra Structure

The current understanding of RecA NPF structure, as reviewed in chapter 3 (section 3.2.2), is derived from a combination of X-ray crystallographic[364,151,149,22], small angle neutron

scattering[208,206,207] and Electron microscopy (figure 9.7 A).[365,19,20,21] These depict that in the presence of ATP or ATPγS, RecA protein forms an active right-handed helical filament with approximately 10 nm width and 9.5 nm pitch with 6 RecA monomers per turn. The encapsulated DNA within the filament is extend by about 50% relative to the standard B-form DNA, with approximately 6.2 RecA monomers (hexameric repeat) and 18 DNA base pairs per turn.

It is important to note that these details are derived from averaged representations, which obscure any localised structural variations. In contrast, single molecule observations of active NPFs have also been demonstrated by both scanning tunnelling microscopy (STM) (figure 9.7 B)[366,367] and AFM using carbon nanotube probes (figure 9.7 C).[368,209] These studies resolve a NPF of uniform width and the helical groove as regular banding perpendicular to the long axis. Interestingly, similar spatial resolutions were achieved with off-the-shelf cantilevers due to a rigorous purification and sample preparation methodology, presented in the author's Masters dissertation.[211] However, all these studies have revealed a compressed and dilated structure, lacking resolution of the fine helical detail when compared to the crystal structures (figure 9.7 A - D).

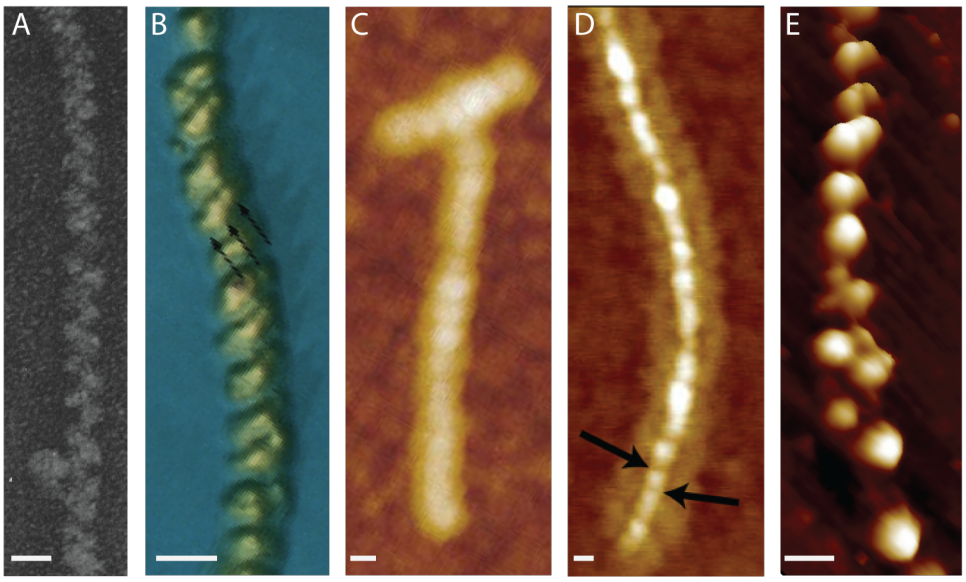

Figure 9.7: ***A comparison of NPF resolution achievements.*** *A series of micrographs depicting a variety of resolution achievements on RecA nucleoprotein filaments taken with:* ***(A)*** *Electron microscopy,* ***(B)*** *Scanning tunnelling microscopy and* ***(C - E)*** *AFM. Of the AFM images:****C)*** *was captured with CNT probes,* ***D)*** *using thorough purification and* ***E)*** *using a rapid force-curve-based imaging technique – observed in this study. Images adapted from: (A) Dunn et al,[365] (B) Amrein et al,[367] (C) Umemura et al[209] (D) and A.J. Lee masters dissertation.[211] All scale bars = 10 nm.*

In this work, NPF formed on 890 bp dsDNA were deposited in 10 mM $Ni^{2+}$and incubated on the surface for 1 hour, as previously described (see section 9.2.1). When imaged with rapid force-curve-based imaging – with a typical estimated load of 100 pN – the helical repeat of RecA NPFs is clearly resolved as a deeply grooved corrugation along the length of the nucleoprotein structure (figure 9.7 E), shown here for filaments formed upon 890 bp dsDNA. From the data presented here, a helical pitch of 10 $\pm$ 1.2 nm with a groove depth of 3.2 $\pm$ 0.5 nm is observed, found to be in good agreement with the previous crystallographic reconstruction[22] (figure 9.7 E).

Interestingly, as previously eluded to by others,[361] the applied load has a large impact on the attainable resolution. Here, applied loads above 200 pN typically resulted in a reduction in the measured height and broadening of the structure consistent with compression of the filament (data not shown) – this was associated with a reduction in access to the helical groove. In contrast, loading below 100 pN was found to result in a reduction in the Z resolution into the helical groove.

Under these conditions, the RecA NPF helical pitch was observed as an alternating curvature between subsequent hexameric repeats, following the linear long axis of the nucleoprotein filament (figure 9.8 A). This structure is found to be consistent with that of the 6 RecA monomer and DNA complex reconstructed by Chen et al (PDB ID: 3CMW) (figure 9.8 B).[22]

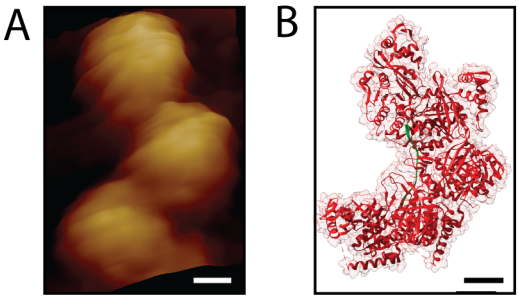

Figure 9.8: ***AFM observation of the RecA nucleoprotein filament hexameric structure.*** *A direct comparison of the AFM* ***(A)*** *and crystal structure* ***(B)*** *of the helical repeat structure of RecA nucleoprotein filaments. Adapted from the crystal structure published by Chen et al. PDB ID: 3CMW*[22]*. Scale bar = 2 nm. Z scale = 10 nm. The molecular model is appropriately scaled for comparison.*

It is important to note that the alternating chirality of opposing banding apparent in the AFM images depicted here, are found to be deep and shallow, respectively (figure 9.9 A & B). Recall that the NPF structure has a relatively large vacancy at its core. When this was taken together with the previous discussion of loading forces, it was postulated that the shallow grooves represent a slight deformation of the hollow structure by the AFM probe where the helical groove crosses the back face of the filament (figure 9.9 C).

This makes sense when it is considered that as the helical groove spirals around the structure it

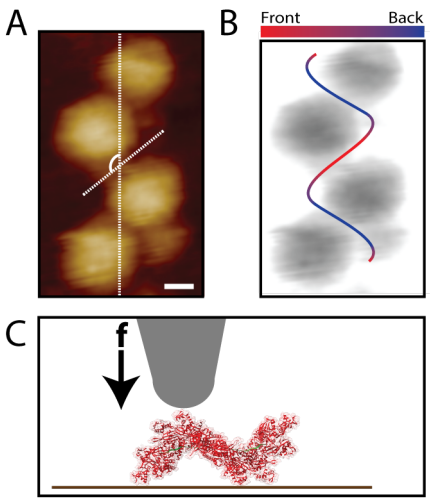

Figure 9.9: ***Structural compression and the NPF helical groove.*** ***A)*** *An AFM imaged depicting a portion of a NPF captured with rapid force-curve-based imaging. The central axis of the filament is marked with a white line, from which the angle of the helical groove is measured.* ***B)*** *The helical groove can be traced around the front (red) and back (blue) faces of the filament structure, evident due to the compression of hollow structure by the AFM probe* ***(C)****. Scale bar = 2 nm. Z scale = 8 nm.*

would cross the back face with the reflected angle (when viewed from the front face) compared to the groove crossing the front of the filament. Further, the groove crossing the back face would be offset by approximately half a full turn, i.e. 5 nm, to that of the observed 10 nm banding across the front face (figure 9.9 A & B).

Furthermore, this phenomena may also give rise to the reduced height, broadened and uniform width of RecA filaments previously observed by others, as a consequence of the hollow structure collapsing on to the supporting substrate. This maybe due to either the sample preparation[366,367] or uncontrolled probe loading,[368] of previous studies. With respect to this, the filaments observed here are found to have a height of 7.2 $\pm$ 0.5 nm, compared to 3.8 $\pm$ 0.5 nm previously reported by Umemura et al.[209]

Moreover, the measured angle of the helical groove as it bisects the long axis of the filament is found to be 124° $\pm$ 2.7° (figure 9.9 A). This is interesting, given the angles of incidence - of 111° $\pm$ 2.7° - measured for NPFs undertaking a homology search in the previous chapter (section 8.2.8). These observations fit well with the proposition of a transient association aligned within binding site II – present within the helical groove (see section 3.2.1) – as suggested by others.[221,160] Given the transient nature of this interaction, it is likely that there is a degree of angular flexibility between the incoming dsDNA and the NPF helical groove, possibly giving rise to the difference in measured angles. Further discrepancies are likely to arise from the reduced spatial resolution of the dynamic observations (section 8.2.8). Nonetheless, when taken together these observations provide the first direct evidence in support of an orientated interaction through the helical groove as predicted by the theoretical work of Klapstein et al[221] and Yang et al.[160]

Discontinuous Nucleoprotein Filament Structure

Surprisingly, this study found discontinuities in the helical structure of the RecA filaments (figure 9.10, orange arrows), not previously reported. It can be argued that this is related to inconsistencies in the polymerisation of the filament – where RecA binds 3 nt ssDNA per monomer (section 3.2.2) – possibly as a consequence of multiple nucleation sites of RecA polymerisation.

Alternatively, these inconsistencies may represent breakage or dislocation of the completed NPF structure upon sample preparation. However, the frequency and consistency with which these dislocations appear across the observed NPF populations is considered to be too high for this to be the case, where the samples are not subjected to any purification or processing prior to observation.

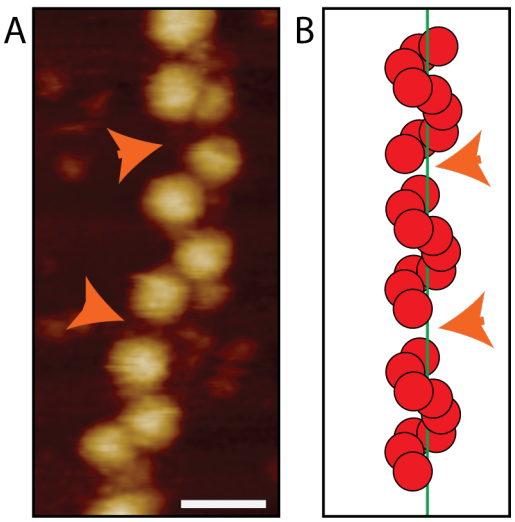

Figure 9.10: ***Observation of discontinuous nucleoprotein filament structure.*** *The structure of RecA nucleoprotein filaments is demonstrated to be less uniform than previously reported, seen in the AFM micrographs* ***(A)*** *and schematically* ***(B)****. Clear discontinuities are highlighted* ***(A), orange arrows****, suggesting polymerisation occurred from multiple locations. Scale bar = 10 nm. Z scale = 10 nm.*

Here, the observed average filament sections, indicated between the orange arrows (figure 9.10), are estimated to consist of 12 - 18 RecA monomers (20 - 25 nm in length). However, it is important to note that due to the limited sample size a statistical argument for the size of individually nucleated filament sections cannot be justified. These observations are nonetheless significant, as it has previously been suggested that it requires nucleation of 5 - 6 RecA monomers to initiate polymerisation.[7] This leads one to conclude that, at least under these experimental conditions, continuous polymerisation accounts for only very short – 12 - 18 monomer – regions of continuous filament growth. These observations highly suggest that filament nucleation occurs from multiple sites, not a single terminus[7] and therefore has implications for the current proposed kinetics of filament growth.[7]

Furthermore, the presence of dislocations has implications for the current understanding of the homology probing mechanism (section 3.2.3). To date, discussions suggested that probing

for homology occurs within short, 6 - 8 nt regions.[6,194,195] When homology is located the synaptic complex becomes stabilised and neighbouring sections undergo sequence probing in an accelerated manner, eventually transitioning to a postsynaptic joint (section 3.2.3).

It is immediately evident that this process will proceed most efficiently where a continuous helical groove enables unimpeded access into the centre of the RecA complex where the encapsulated ssDNA bases are exposed. Thus, the observation of dislocations in the NPF structure raises some interesting questions regarding the efficient transition from synaptic to postsynaptic joint. How does RecA account for these inconsistencies during sequence probing? What happens if two regions are probed simultaneously within different RecA NPF regions? Can this lead to kinetically trapped complexes? It is plausible that the presence of dislocations is tolerated by RecA, where evidence suggests that native complexes formed with ATP may be able to rearrange to close or accommodate discontinuities in the polymer structure. This is due to the continual turn over of the nucleoside co-factor at the level of individual monomers – compared to cooperatively throughout the NPF during de-polymerisation in the postsynaptic phase – which would suggest that NPFs are considerably more dynamic than typically described.[152]

Evidently, direct single molecule visualisation offers potentially insightful quantification of local structural irregularities that are unavailable to ensemble methods. One such example where this is true, is the interrogation of local variations in the helical pitch of dsDNA as a consequence of its manipulation by RecA. Hence, the methodologies that have enabled the structural observations described thus far in this chapter are applied to this investigation, this is elaborated on in the next section.

9.2.4 Observing the RecA Induced Helical Compression of DNA

To identify the structural manipulation of DNA by RecA, any induced variation must be isolated and prevented from dissipating. As suggested (section 9.1), in order to trap any helical compression induced by RecA the termini of the subject DNA molecules must be anchored to restrain their rotational freedom. Prior to the introduction of NPFs, efforts were made to establish the helical pitch of dsDNA central strands.

In the previous chapter (section 8.2.1) DNA origami frames were utilised to anchor subject DNA molecules for the real-time observation of NPF interactions with DNA. In these experiments, the rotational freedom of the DNA molecules was preserved by anchoring into the origami support frames using a single strand.

The DF2S frame was therefore re-used for the experiments in this section, where the anchor points were modified to bind both strands - upper and lower - of the DNA molecule into the frame at either termini (figure 9.11). This acts to restrict the rotational freedom of the DNA molecule, trapping any induced strain as a result of the subsequent RecA interaction. The 64 b central DNA strands were annealed from separate oligonucleotides, with 16 b overhangs complementary to the upper and lower portions of the frame anchors, as appropriate (figure 9.11). These central strands were found to incorporate into DF2S with at least 90% efficiency at a 5:1 (DNA:Frame) ratio; with the excess removed through subsequent purification (see methods 5.2.9).

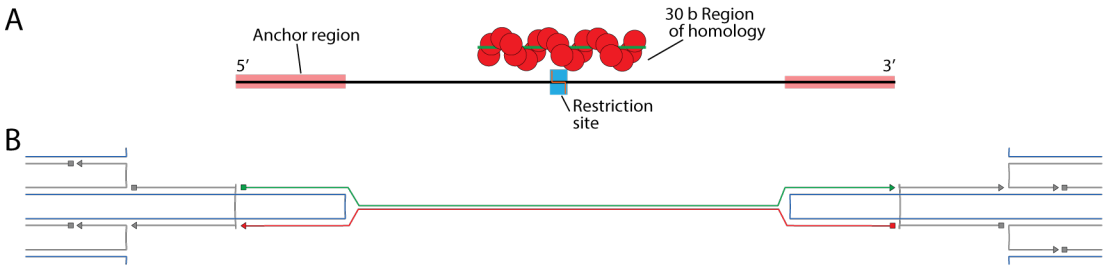

Figure 9.11: ***A schematic diagram depicting the strand arrangement for viewing the RecA induced helical compression of DNA. A)*** *The arrangement of features within the central strand incorporated for DF2S. The 30 nt NPF is patterned over the central restriction site. The anchor regions are highlighted in red.* ***B)*** *The double anchor layout utilised to remove the rotational freedom of the dsDNA molecule is depicted, enabling the trapping of induced helical compression.*

Conditions for the strong surface adhesion of the central strands – sufficient for obtaining helical pitch resolution – was investigated prior to the introduction of NPFs. Consistent resolution of the helical pitch on these central DNA strands is critical for detecting variation where RecA is introduced in subsequent experiments. Initially deposition was carried out as had proven successful for both the linear DNA (see section 9.2.1) and NPFs, with the inclusion of Ni^{2+}in the deposition buffer. However this was quickly found to be inappropriate here as this tended to cause aggregation of the origami frames, leading to large surface clusters. In contrast, where surfaces were pre-incubated with 1 - 10 mM Ni^{2+}- and deposition was carried out in the origami folding buffer, containing 10 mM Mg^{2+}- the appearance of rolled up origami frames predominated. This is most likely a consequence of inherent warp in the frame.

Alternatively, where the frames were deposited only in the presence of $Mg^{2+}$and allowed to equilibrate with the surface for up to 1 hour, excellent yields were achieved. However, despite these conditions being conducive to decent resolution on the origami frame itself, the central strands remained difficult to resolve (figure 9.12 A).

From figure 9.12, the flexibility of the central strands can clearly be seen, owing to the inability of the scanning probe to resolve them. This is strongly indicative of the loose surface association of the central strands, poorly resisting the forces of the scanning probe. Although this is the largely desired characteristic for the previous dynamic observations, it is highly counter-productive to the experiments conducted in this section.

In response to this, it was concluded that the introduction of Ni^{2+}following the successful deposition of the origami frames would help to bind the central strands, without encouraging negative aggregation or rolling of the frames. A series of post-deposition Ni^{2+}incubations from 1 - 10 mM and from 5 minutes to 2 hours were conducted, all of which indicated an enhancement in the central strand surface binding. In particular, the introduction of 10 mM $Ni^{2+}$was found sufficient to orchestrate successful binding characteristics after 5 minutes (figure 9.12 B, with Nickel). Longer incubations or higher concentrations typically were found to cause unwanted crystallisation effects. However, despite this subsequent efforts to resolve the helical pitch of these central strands remained unsuccessful (figure 9.12), where the non-conducive mobility was still evident.

This problem can be reconciled when the frame itself is re-examined, where finite-element-based simulations with single bp resolutions – conducted using the online resource CanDo[65,369] – highlighted the presence of a severe global curvature in the frame (figure 9.12, bottom). In the first instance, this provides a likely cause for the observed rolled filamentous origami structures when deposited in Ni^{2+}. Furthermore, as it has already been demonstrated that internal strain or warp persists in DNA origami tiles when deposited upon mica substrates (section 9.2.2), then this may go some way to explaining the apparent persistent flexibility of the central strands. One may consider that any built up strain within the frame would curve the frame towards its weakest point, in this case the large vacancy at its centre - this can clearly be seen in the molecular dynamics simulations (figure 9.12, bottom). This effectively acts to push the anchors and hence the central strands up and away from the surface, rendering them extremely flexible - an effect that evidently persists despite the localisation of Ni^{2+}.

In light of this limitation and when taken together with the difficulties experienced when attempted to achieve the introduction of only a single 30 nt NPF within each frame, it was largely concluded that this work couldn't proceed any further without a major experimental redesign. This required the complete re-development of the DNA origami frame to neutralise the internal strain and the development of a high yield purification protocol that could be conducted following the introduction of NPFs at much higher ratios than conducted in the experiments in this thesis. Given

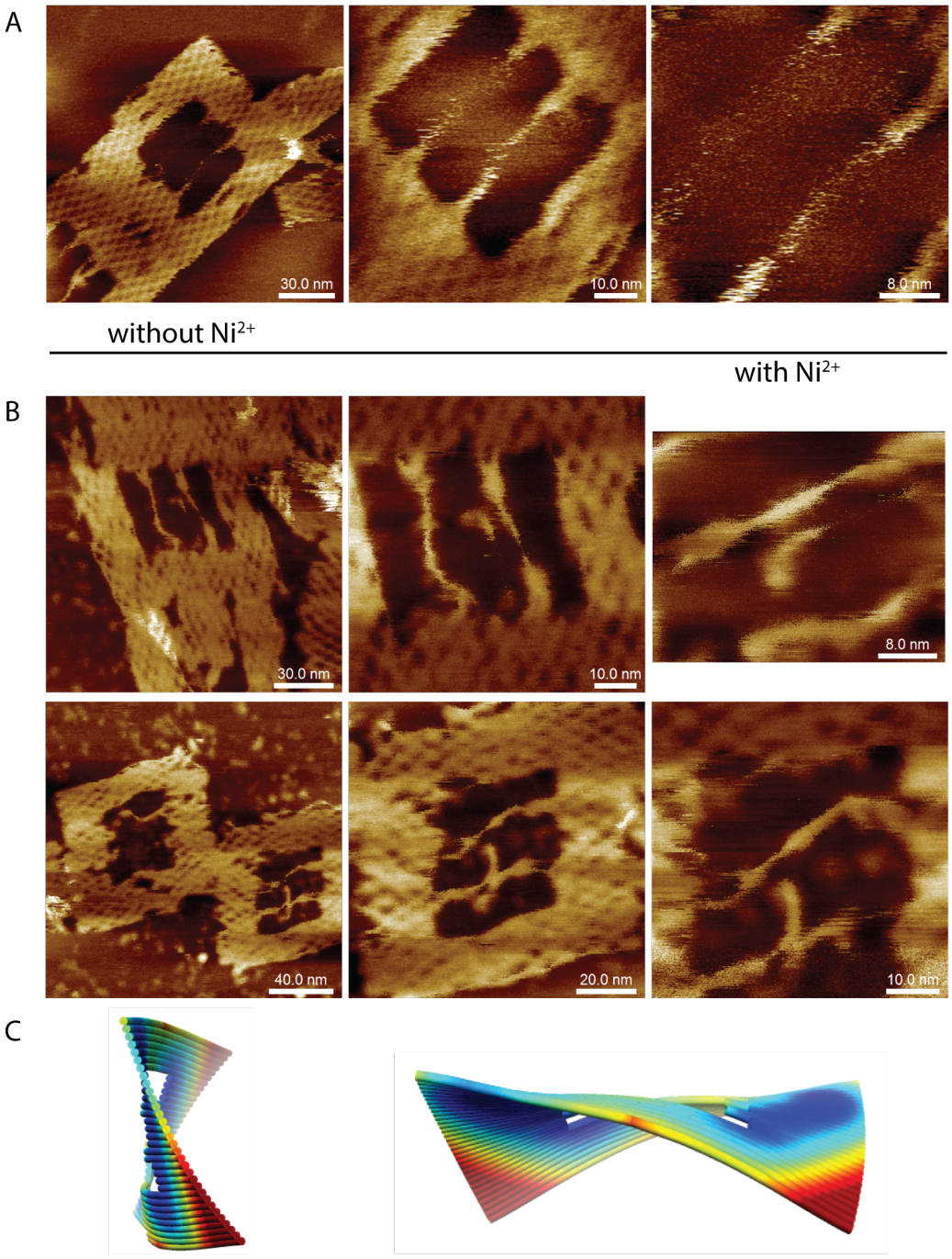

Figure 9.12: ***A montage of AFM images depicting the attempts to image the helical pitch of the trapped central strands.*** ***A)*** *The central strands are seen to be highly flexible in the absence of* Ni^{2+}. ***B)*** *In contrast, the central strands were bound tightly when* Ni^{2+} *was introduced post deposition of the DNA frame. However the resolution remains limited.* ***C)*** *A finite-element-based simulation conducted with single bp resolution reveals a large internal strain present within the frame. Simulation undertaken using the online resource CanDo.*[65,369]

the time constraints of the project, this was not undertaken and hence future work is required.

9.3 Conclusions

In this chapter, rapid force-curve-based AFM imaging has been successfully applied to the observation of RecA NPFs and complex DNA topologies, revealing novel structural details for the first time. This work was undertaken with the ultimate aim of examining the structural manipulation of DNA by RecA.

Firstly, observations of the DNA helical pitch has been explored and expanded to complex DNA topologies, demonstrated here for the first time throughout DNA origami. This has significant implications for the future scrutiny and assessment of individual DNA structures or specific regions therein. Enabling direct characterisation of localised internal strains, porosity of the weave and longer scale warp or periodicity within the structure. Such detailed information is important when developing novel devices from any material and is hence likely to aid the development of DNA structures and the inclusion of DNA binding proteins, specifically RecA NPFs.

Furthermore, examination of RecA nucleoprotein filaments in isolation has revealed the helical hexameric subunit, complementing the previously described crystallographic reconstructions.[22] Interestingly, this superior resolution allows the identification of heterogeneity within the helical structure, previously unobserved by other methods, thus providing direct evidence for multiple nucleation sites during RecA polymerisation. However, as discussed, due to the small sample size in this work, conclusive statistical measurements of RecA nucleation sites cannot be made, making this an area for future investigation. Such detail may have a crucial impact on our understanding of the RecA homologous strand exchange mechanism and how the impact of discontinuities effects this interaction over long sequences.

Unfortunately, the overall goal of this chapter remains unaccomplished by the time of writing as a result of limitations in the experimental design. It was found that an inherent warp in the DNA origami support structure was the likely cause which prevented the sufficiently strong binding of the subject strands for high resolution observations. In order to rectify this, a drastic re-design of the support structure is required to render the internal strain neutral, which sadly due to time constraints has not been realised. However, despite this, the work in this chapter has demonstrated several significant resolution achievements, surpassing those reported in the literature.

In conclusion, this work demonstrates that AFM, in particular rapid force-curve-based imaging

methods have a critical role to play in describing the local structural variations within isolated biological molecules and derived structures, providing deeper insight into their inherent functions.

Chapter 10

Conclusions and Future Perspectives

10.1 Conclusions

The formation of nanoscale functional devices from biological molecules may be considered the ultimate goal for bionanotechnology. These bottom-up approaches based on molecular self-assembly have the potential to surpass the most advanced top-down microfabrication techniques, which are reaching the physical limits of the current technology.

It can be argued that DNA nanotechnology presents the most promising route to realising the rational design of devices from biological molecules. Over the last three decades this field has seen rapid growth from proof of principle to practical application. Arguably the most critical advance has been the development of scaffolded designs, enabling woven origami structures or large arbitrary structures formed from wire-frame polyhedra.[60] Which demonstrates that arbitrary patterns can be formed in both two and three dimensions.

Moreover, recent advances in organised surface deposition hails in an era where such solution based systems may be integrated directly with micro-fabricated structures to interface directly with the macroscale world. Hence, one may consider that DNA nanotechnology acts as an adapter for templating functional components - such as transistors - which may be interfaced together to form larger circuits and devices.

Considering DNA origami as a molecular "bread-board" in this way, it is evident that efficient methodologies for orchestrating the placement of components is required. One approach, being developed in the Bioelectronics group, is to harness the homologous recombination mechanism of the RecA protein. RecA mediates the formation of a triple–stranded DNA complex where sequence homology is shared between ssDNA and dsDNA species. Harnessing the homologous

recombination mechanism of RecA enables the addressing of a DNA substrate arbitrarily and sequence specifically, without the need to redesign the underlying template. For example, a single DNA origami tile design may be addressed in many different ways to form a multitude of components, enabling the tile design to focus on form factor for optimal surface placement and integration between neighbours.

The applicability of RecA has already been demonstrated by the work of Keren et al[9,10] and Nishinaka et al[8]. Further work in the Bioelectronics group has demonstrated that this approach can be successfully multiplexed and scaled down to allow feature sizes of less than 10 nm.[4,5,6]

However, despite the success of these studies, severe gaps remain in understanding the fundamental biological mechanism that underpins this approach. As such, the exact mechanism by which NPFs are able to locate regions of homology within a DNA template remains widely debated. Not only is it critical to answer this question from a biological context, but it has direct implications for any bionanotechnological application of RecA, such as its interaction within artificial complex DNA topologies. In response to this challenge, the work presented in this thesis has applied recent advances in spatial and temporal resolutions of the AFM to directly interrogate the homology searching mechanism of RecA NPFs.

The aims of this project were to directly observe the NPF homology search on dsDNA and assess its primary mode of interaction; one dimensional facilitated diffusion or random contact (three dimensional diffusion). In addition to exploring the structural - and hence mechanical - manipulations of dsDNA by NPFs which may limit the ability to affect efficient RecA mediated patterning within constrained artificial DNA complexes.

From the literature published to date, the homologous recombination mechanism of RecA has been investigated in broadly two ways. Those studies which attempt to imply function from structure and those studies which attempt to directly interrogate function.

To date, crystallographic structures which pertain to different intermediates of the mechanism have been presented.[21,12] However, differences in the structures between studies suggest a degree of sample preparation artefact, undermining the designated status of each structure as distinct steps in the process. These structures have further informed molecular dynamics simulation approaches which begin to suggest that transient associations of DNA with secondary binding sites forms the bulk of synaptic contacts.[161,221,160]

In contrast, dynamic studies have demonstrated that the three dimensional degrees of freedom of the dsDNA molecule is of critical importance, the degree to which heterology can be overcome,

along with broadly speculative rate constants in addition to providing evidence in support of transient binding intermediates. However, all of the single molecule approaches employed to date either use indirect reporting - for example FRET[15,183] - or lack the resolution - in the case of fluorescence measurements[11] - to discern critical parameters such as the number of interacting species, the orientation of interactions or the prominent mechanism by which NPFs sample the sequence space of the DNA substrate.

The work presented in this thesis, in contrast, demonstrated direct observation of the mechanisms by which RecA NPFs conduct a search for sequence homology (chapters 7 & 8). This work has provided direct evidence for the existence of random three dimensional contacts or hoping between locations on the dsDNA (chapters 7 & 8). Furthermore, this is accompanied by a one dimensional diffusion along the DNA substrate (chapter 7) which is observed to occur at a characteristic angle (chapter 8). This is suggestive of the DNA travelling within binding site II aligned along the helical groove of the NPF as demonstrated by the high resolution observations made in this study (chapter 9).[22,160] Moreover, NPFs are shown to form associated clusters, locally enhancing their concentration in the vicinity of the DNA substrate (chapters 6 & 7). This is likely to increase the homology searching efficiency as a result of parallel interactions occurring simultaneously (chapter 6).

10.1.1 Cooperative RecA Interactions Increase Homology Searching Efficiency

The first approach taken by the author in this work was to evaluate the nature of intermediates across the typical time scale of a RecA patterning reaction (chapter 6). Work within the Bioelectronics group had demonstrated high patterning efficiencies with increasing ratios with a standard incubation of 60 minutes. Homology searching interactions conducted in solution and at a solid-liquid interface were ceased at specific time points over the course of a 60 minute interaction and the resultant populations of intermediates were statistically assessed with the AFM.

This work revealed that the interaction landscape is considerably more complex than demonstrated to date, occurring in a massively parallel manner. NPFs were shown to interact with dsDNA in a cooperative manner resulting in the rapid 'association' of multiple synaptic joints per DNA molecule. In this work, 890 bp DNA molecules were found to have an average of 3 synaptic complexes, with up to 7 observed on the template simultaneously. This represents nearly half of the entire sequence space being sampled simultaneously, hence it is easy to see how this increases the homology search efficiency.

Moreover, these associations are found to be short-lived where homology is not located, indicating that they occur through transient associations. Such heterologous joints are observed to undergo a 'resolution' phase, resulting in a singular homologous pairing per template. Notably, the resolution of heterologous synaptic joints was found to be hindered by the presence of a solid-liquid interface.

As the corresponding association phase was only partially perturbed it is suggested that it is the continued reduction in degrees of freedom of the DNA template - with successive NPF complexes binding - that limits this process. Indeed, characteristic 'bunches' of NPFs were observed side-by-side on DNA templates.

This not only suggests that the DNA is stabilised within the complexes, but also demonstrates that NPFs can co-exist directly next to one another. This is interesting as it is known that the dsDNA is under-wound within the complex, hence little or no transmission of this mechanical strain is suggested between neighbouring complexes which may limit neighbouring synaptic joint interaction.

It must be considered that even though the DNA templates in question are bound at a solid-liquid interface, they are likely to largely retain rotational degrees of freedom and hence any induced mechanical strain caused by the under-winding within an NPF is likely to dissipate through rotation of the molecule.

In contrast, where the geometry is fixed, for example in between branched junctions within a synthetic woven DNA structure, such rotational freedom is restrained. Hence, it remains to be seen whether such close association of NPFs is possible in complex templates. Attempts to explore this matter were conducted in chapter 9 and are discussed in the proceeding sections.

The occurrence of parallel multiple-synaptic joint formation is reconciled by the observation of the formation of RecA nucleoprotein clusters. These are resolved throughout these static studies suggesting that the cooperation between NPFs acts to increase the local concentration in the vicinity of the DNA scaffold, thus facilitating a more efficient interaction through the confinement of reactive species.

Moreover, the observation of these clusters in dynamic observations with HS-AFM demonstrate that the NPFs are loosely associated and that they can appear directly from solution. As such, this suggests that they do not represent detrimental aggregates, but remain in an active state. Interestingly, it is known that RecA can undergo inter-filament interactions, typically observed as bundles of long filaments twisted around one another.[322] Furthermore, recent work has suggested that these bundles are active elements of the recombination mechanism *in vivo* aiding in the pairing

of distant genomic loci.[323] Hence, the appearance of the clusters observed in this work may derive from similar origins.

Regardless of the mechanism, it became immediately apparent that the cooperative nature of NPFs may present difficulties when patterning is applied to complex or confined DNA geometries. Furthermore, directly concerning this work, it was found to be a point of contention where singular NPF:DNA interactions were to be observed with the HS-AFM.

10.1.2 Demonstrated Control Over the Translational Freedom of DNA and Molecule Tracking

In order to directly observe the homology searching mechanism of NPFs a sample preparation was developed to enable a transient absorption of dsDNA molecules at a mica surface.

Here, the subject DNA molecules are required to be associated with the surface strongly enough to enable stable imaging over relevant time-scales, while at the same time providing enough translational freedom to allow the proteins to interact with the DNA in a way that is as undisturbed as possible.

The approach presented in this thesis successfully tuned the surface interaction of DNA, by leveraging the different binding capacities of $Mg^{2+}$and Ni^{2+}divalent cations. Here the concentration of Ni^{2+}utilised for pre-incubation of the mica surface acts to restrain the DNA on the surface. This is balanced with a constant 10 mM Mg^{2+}in the imaging buffer which mediates weak DNA-surface association and acts as the reaction co-factor for subsequent nucleoprotein interactions. Hence the translational freedom of the surface bound DNA was tuned as a function of the concentration of Ni^{2+}mediated strong binding sites distributed across the surface.

In order to assess the translational freedom orchestrated by this approach an algorithm for tracking DNA molecules was developed. When taken together, these approaches enable the appropriate tuning of different length DNA molecules suitable for a wide range of nucleoprotein interactions.[320]

Further to this, the presented analytical approach enables the analysis of the impact of various conditions to the subject DNA molecules, including the interaction of the AFM tip and the existence of specific surface pinning points.

These approaches were successfully applied to the observation of RecA dynamics, in particular the homology searching behaviours of NPFs on dsDNA. These direct observations provide clear

evidence for the existence of a facilitated diffusion along the DNA template, previously disputed by others. Moreover, the prevalence of cooperative action is demonstrated through the observation of NPF clusters, providing support of their existence as observed from the previous work.

However despite these successes, due to the inherent mobility of both molecules (DNA and NPF) orchestrated by this sample preparation method, it remained difficult to reconcile specific quantities of the interactions desired by this study. In particular, it is difficult to infer any directionality to the observed interactions, without the additional inclusion of a polarity marker to the DNA template. Furthermore, quantifying the distance and hence velocities of NPF sliding is inhibited by the lack of fixed references from which to measure. These limitations in the experimental setup proved difficult to overcome and hence alternative approaches were investigated. Nonetheless, the approach developed here remains valid for the observation of large scale manipulations of DNA molecules such as the action of restriction enzymes - as demonstrated here for EcoRI - or the action of nucleosome wrapping, for example.

10.1.3 RecA NPFs Shown to Undertake Orientated Sliding Over Short Distances When Observed in Origami Reference Frames

In order to circumnavigate the experimental limitations of the previous method, the DNA origami frame approach of Sugiyama and Endo et al[128,236] was adopted. This methodology utilised DNA origami as a reference from which orientation and position of the interaction species was quantified.

In the first instance, a design was re-created from the literature and was successfully utilised to observe the occurrence of NPF hoping and sliding while conducting a homology search. Limitations in the sequence length available within this design was rectified through the collaborative development of a larger DNA frame with Prof. Masayuki Endo (Kyoto University, Kyoto).

Interestingly, from these observations NPFs are seen to diffuse an average of 10 nt along the dsDNA molecule. This is approximately 6 times shorter than those suggested by Ragunathan et al using FRET studies in solution.[15] This difference is reconciled by the presence of the solid-liquid interface in the work presented here. Moreover, despite demonstrating that processive movements of the NPFs are largely independent of the orientation of the scanning probe, its interruption of the interaction cannot be ruled out. Hence, it must be considered that the instantaneous interactions of the scanning probe with the subject NPFs may disrupt the transient interactions with the DNA,

limiting the length of stable contacts required for sliding to occur. As such, the extraction of interaction rates from these experiments must be considered with care and are likely not to be representative of a physiological scenario, measured to be 5.73 bp s^{-1} $\pm$ 0.47 bp and 1.1 bp s^{-1} $\pm$ 0.18 bp, for hops and slides, respectively.

Nonetheless, the direct observation of sliding in this work remains significant as it is the first direct confirmation of the existence of this process. Furthermore and most importantly, the work presented here demonstrates that NPFs interact in a specific orientation with the dsDNA molecules. Such detail supports the molecular dynamics simulations conducted by others,[161,221] in particular Yang et al.[160] Here, demonstrating for the first time experimentally that NPFs interact with dsDNA in a specific orientation consistent with alignment within the negatively charged tract of the helical groove – as suggested by Yang et al.[160] Here, a typical interaction angle of 111° is observed for NPFs observed during dynamic observations with the HS-AFM. These measurements are reasonably well complemented by the high resolution observations made in chapter 9, demonstrating helical groove angles of 124°, with respect to the central axis of the NPF. Given the discrepancies in these measured angles, further work is required to interrogate the true significance of orientation upon the NPF:DNA interaction.

Despite the clear observation of transient homology probing interactions, limited evidence for the transition of a pre-synaptic to a synaptic complex is evidenced. Nonetheless, successful patterning was demonstrated for the first time in complex DNA structures. This is critical in achieving the desired goal of patterning arbitrary locations within complex DNA architectures, demonstrating that NPFs can interact with more complex geometries. Attempts were further made to interrogate the effects of mechanical manipulation of DNA within complex geometries with limited success (chapter 9).

Typically, successful patterning was not directly captured – but evidenced in its terminal state – highlighting that despite excellent HS-AFM experimental setup, one must observe in the right place at the right time. This is likely to be a function of the stringent concentrations utilised for these experiments. As suggested, the cooperative nature of NPFs presents a problem for observing singular interactions. Hence, for this work, low concentrations were utilised to minimise the formation of clusters which negatively impacts on the populations of interacting species observed.

Furthermore, it must be considered that the presence of the solid-liquid interface is likely to play a large roll in limiting the full wrapping of a target DNA molecule into the NPF complex. Although it was noted in chapter 6 that the surface only posed a limitation to the resolution of interacting

NPFs. It must be considered that experiments were conducted on free linear DNA molecules, able to increase the radius of gyration locally by reducing the separation between the molecules termini, which can typically be seen in DNA condensates.

In contrast, where the termini of the DNA molecules are tethered within an origami frame structure, their position is fixed and the molecule may therefore be considered to exist under constant tension. Hence, the radi of gyration remains relatively constant, likely insufficient to wrap into the NPF complex. This is supported by the fact that NPFs pattern successfully into the frames with increased efficiencies in solution compared to the surface - suggesting that here the NPF itself is able to rotate about the DNA molecule to overcome its limited gyration. This however, is severely hindered at a solid-liquid interface.

Potentially, an increase in flexibility can be accommodated within the larger DNA frame developed in collaboration with Prof. Masayuki Endo. However it was found that stretching the geometry of the frame to this size whilst still using the same M13mp18 template, resulted in detrimental loss of structural integrity. Due to the time restrictions within this study, further development has thus far not been pursued (see chapter 9). It is proposed that going forward, a larger ssDNA scaffold would be utilised to overcome these problems. Although this approach is uncommon, it has been demonstrated that the origami approach is inherently scalable provided a sufficient quantity of ssDNA scaffold can be generated.[72]

10.1.4 Surpassed Resolution Achievements upon Nucleoprotein Filaments and Complex DNA Topologies

Finally, the structural manipulation of dsDNA by NPFs was investigated in order to understand the implications of attempting to pattern into the constrained DNA topologies present within DNA nanotechnology. This work utilised recent advances in rapid force-curve-based AFM imaging enabling the probe loading forces to be maintained at a minimum – typically 100 pN used in this work – giving access to superior resolution on soft matter.

Initially, the sample preparation described by Pyne et al[361] was utilised to obtain helical pitch resolution of λ phage DNA deposited upon mica. Here, not only were the major and minor grooves of the B-form dsDNA helix evident, but larger scale periodicity was observed representing supercoiling within the molecule.

Furthermore, this approach was applied to DNA origami revealing the helical pitch within this complex woven topology for the first time. Interestingly, the longer scale periodicity was observed

within these structures as well, evidencing the internal strain that results from the non-integer positioning of cross-over points.

This has significant implications for the future scrutiny and assessment of individual DNA structures or specific regions therein; enabling direct characterisation of localised internal strains, porosity of the weave and longer scale warp or periodicity within the structure. Such detailed information is important when developing novel devices from any material and is hence likely to inform the development of DNA structures for specific applications in particular additional components - such as RecA NPFs - are included for hierarchical construction.

Moreover, this work further demonstrated superior spatial resolutions on NPFs observed in isolation, revealing the helical hexameric subunit with AFM for the first time. This resolved structure fits remarkably well with the previously described crystallographic reconstructions produced by Chen et al.[22]

Interestingly, this superior spatial resolution allows the identification of discontinuities within the helical structure, previously unobserved by other ensemble methods. These 890 bp NPFs were found to be formed from several distinct regions of continuous helical geometry suggesting that they were formed from multiple nucleation sites. These discontinuities are likely to have a critical impact upon the ability of the NPF to convert from a synaptic to a postsynpatic complex in a processive manner – particularly where it is suggested that distal portions of the same NPF can interact with different locations on a DNA template simultaneously.[11] Hence, how RecA accommodates these discontinuities is likely to be an area for further investigation.

In addition, it has been suggested that NPFs scan undergo rearrangement of the individual monomers within the complex as ATP is turned over.[7] Such arrangements could act to close discontinuities in the helical groove that derived from multi-site polymerisation. However, it is important to note that the NPFs presented in this work are formed with ATPγS and hence are restricted from such rearrangements. Furthermore, there is some evidence to suggest that additional proteins function to recruit RecA and aid in its polymerisation on DNA. In particular, increased rates of polymerisation are noted for NPFs formed in the presence of single stranded binding protein (SSB), which is shown to remove detrimental secondary structure from ssDNA and aid the loading of RecA monomers. Hence, it is likely that NPFs formed in the presence of SSB will contain considerably fewer irregularities. Although this is important when considering the physiological condition, any bionanotechnology application is likely to be conducted using a minimal reaction setup - hence the presence of such discontinuities must be considered here.

Unfortunately, the final goal of directly observing the manipulation of the dsDNA helical pitch due to the interaction of an NPF remains unaccomplished due to challenges with the sample preparations that could not be reconciled in the time available.

These experiments aimed to trap any induced helical compression in the DNA molecule proximal to the NPF by anchoring the DNA within a DNA origami frame. Here, unlike those utilised for dynamic observations, the subject DNA molecule was anchored with both strands of the duplex, hindering its rotational freedom.

However, it was quickly discovered that an inherent warp in the DNA origami structure - evidenced in the high resolution observations of the structure - prevented the central DNA strand from binding to the mica surface with sufficient strength. Hence, the desired resolution could not be achieved on these samples. In order to rectify this, a drastic re-design of the support structure is required to render the internal strain neutral, which is an area for future work.

10.2 Future Work and Perspectives

Despite several limitations, the work presented in this thesis has demonstrated many novel observations of significant impact. However, there remain many open questions that need to be addressed. This section discusses the loose ends that are to be addressed in the short term and the additional project directions that maybe tackled in the long term.

10.2.1 Immediate Term Objectives

Re-development of DF2L from Longer Scaffold

In the first instance, the redevelopment of the DF2L origami frame is suggested to circumnavigate the limited structural integrity of the current design. As the current design stretches the M13mp18 scaffold to its maximum limits, the use of a non-standard larger scaffold of 8.6 kbp is suggested. This scaffold is commercially available as full M13 bacteriophage for phage display applications in molecular biology (New England Biolabs, MA, USA). Also note that using additional sub-cloning methods from molecular biology, the scaffold could be potentially increased further. Initial work would be required to grow sufficient quantities of this phage and extract the ssDNA for use in origami applications. The enlargement of the scaffold sequence would enable the DF2L frame

to have its structure increased to provide sufficient stability whilst maintaining the 224 bp central space for the continuation of NPF homology search observations across larger sequence space.

Development of a Suitable Frame for Trapping Helical Compression

In order to resolve the helical compression induced by the manipulation of dsDNA by NPFs a subject molecule was anchored into a DNA frame restricting its rotational freedom. The use of the DF2S frame for this purpose was found to be flawed due to internal strain within the structure, preventing the central strands from binding the mica surface sufficiently strongly.

In order to rectify this problem, a new DNA frame design is proposed - largely the same as DF2S - but with extra care taken to neutralise any global warp in the structure. Further considerations of the probe geometry must be taken into account in order to image features on the dsDNA juxtaposed with the much larger diameter of the NPF. Interest in developing probes for such purposes has been expressed by collaborators at Bruker (Santa Barbara, CA, USA), including high aspect ratio and angle corrected probes. It is proposed that interaction with these collaborators would enable the desired resolution to be achieved and would present the first direct observations of protein manipulation of the structure of DNA.

10.2.2 Additional Project Directions

The work presented in this thesis, and the work conducted by others[370] have demonstrate the utility and versatility of the use of DNA origami structures for investigating nucleoprotein interactions. In addition to the observation of the RecA mediated homology search conducted in this thesis, there are several stages in the homologous recombination process that are of interest. Specifically these are the completion of strand exchange to form a new heteroduplex and the polymerisation of monomeric RecA on DNA. These proposed directions are presented below.

Observation of RecA Mediate Strand Exchange

Although the observation of paired intermediates have been observed, to date no real–time observation of a full strand exchange has been demonstrated. However, similar recombination events have successfully been observed by Endo et al, hosted within DNA origami frames - in particular that of the site specific recombinase, Cre.[236]

It is proposed that a three–stranded DNA frame would be used to harbour a central reaction strand, with a region of ssDNA that can be formed into a NPF. Either side of which a dsDNA strand would be incorporated – one heterologous as a control and one sharing homology with the NPF enabling strand exchange to be observed. RecA would be selectively polymerised onto the ssDNA region only and the resultant NPF allowed to interact upon release by the breakage of a photocleavable linker. As the NPF remains tethered to the rest of the frame by the remaining dsDNA portion of the strand, it is envisioned that any resulting complex would be observed as a J structure formed by the resultant interaction between the three DNA strands.

Initial enabling work has already been conducted for these experiments, including; the design of an appropriate DNA frame containing three DNA strands (developed in association with Prof. Masayuki Endo); inclusion of partially single stranded DNA sequences into the frame; and the successful selective polymerisation of RecA on ssDNA have all been achieved.

Investigations of RecA Polymerisation

Further work proposed would be to investigate the polymerisation of RecA on a variety of DNA constructs, including; ssDNA, dsDNA and complex junctures. From the work investigated in this thesis, discontinuities are evidenced in the polymerisation of NPFs suggesting that polymerisation occurs from multiple locations. Furthermore, initial experiments demonstrated the observation of nucleation clusters in dynamic experiments, however full polymerisation was not realised with this sample preparation.

In order to further these observations, it is proposed that polymerisation would be investigated within an origami frame, potentially DF2L. As demonstrated in chapter 7 this is a considerably difficult experiment to conduct considering the large concentrations of monomeric RecA required to reach the critical nucleation concentration in the vicinity of the DNA and hence the surface which typically leads to detrimental contamination. However, it might be possible to polymerise at lower concentrations using the supporting protein SSB.

These dynamic experiments, coupled with associated rapid force-curve-based imaging could enable the investigation of polymerisation from multiple locations. Where experiments would compare the use of SSB, ATP and ATPγS on unique variations within the NPF structure. In addition, differential rates of polymerisation – and dissociation – on ssDNA and dsDNA and how the structure of RecA is organised around complex junctions could be investigated.

10.2.3 Perspective on the Field

The work presented in this thesis and the seminal work by others[270,370,361] has demonstrated the applicability of investigating biological molecules with the AFM, informing their application in bionanotechnology. Furthermore, the use of bionanotechnologically derived structures – such as DNA origami – critically enhances the experimental tool set for isolating and interrogating specific interactions at these length scales. Given the desire to realise practical applications from DNA nanotechnology it is clear that such AFM investigations have a large role to play in supporting these developments.

Considering DNA origami as a molecular "bread-board", as previously described, it is evident that efficient methodologies – such as the RecA base approaches being developed by the Bioelectronics group – for orchestrating the placement of components is required. Work to date has demonstrated the scalability of this approach, with $<$10 nm features achievable.[6] However, practical realisation of patterning these features within a woven DNA structure is yet to be demonstrated. Furthermore, investigation of the homologues of RecA, such as Rad51, RadA and Dmc1 may offer improved efficiencies or further available functionalities to be orchestrated.

As such the programmable specificities of these proteins, in conjunction with those whose binding is determined by pre-defined sequences, DNA structure – such as within the major groove – or specifically at junctions between DNA molecules may enable the development of a unique toolbox for the hierarchical spatial addressing of DNA structures. As such, the investigation methods utilised in this thesis may therefore be successfully applied to a large variety nucleoprotein interactions in an attempt to examine the structural variation which may ultimately impact upon the complexes function or application in non-native environments.

This page is intentionally left blank.

Appendices

Appendix A

DNA Origami Frame Schematics

In this appendix the cadnano designs for DF2S and the three iterations of DF2L are given. The M13mp18 bacteriophage ssDNA strand is depicted in blue and the staple oligonucleotides are depicted in various colours.

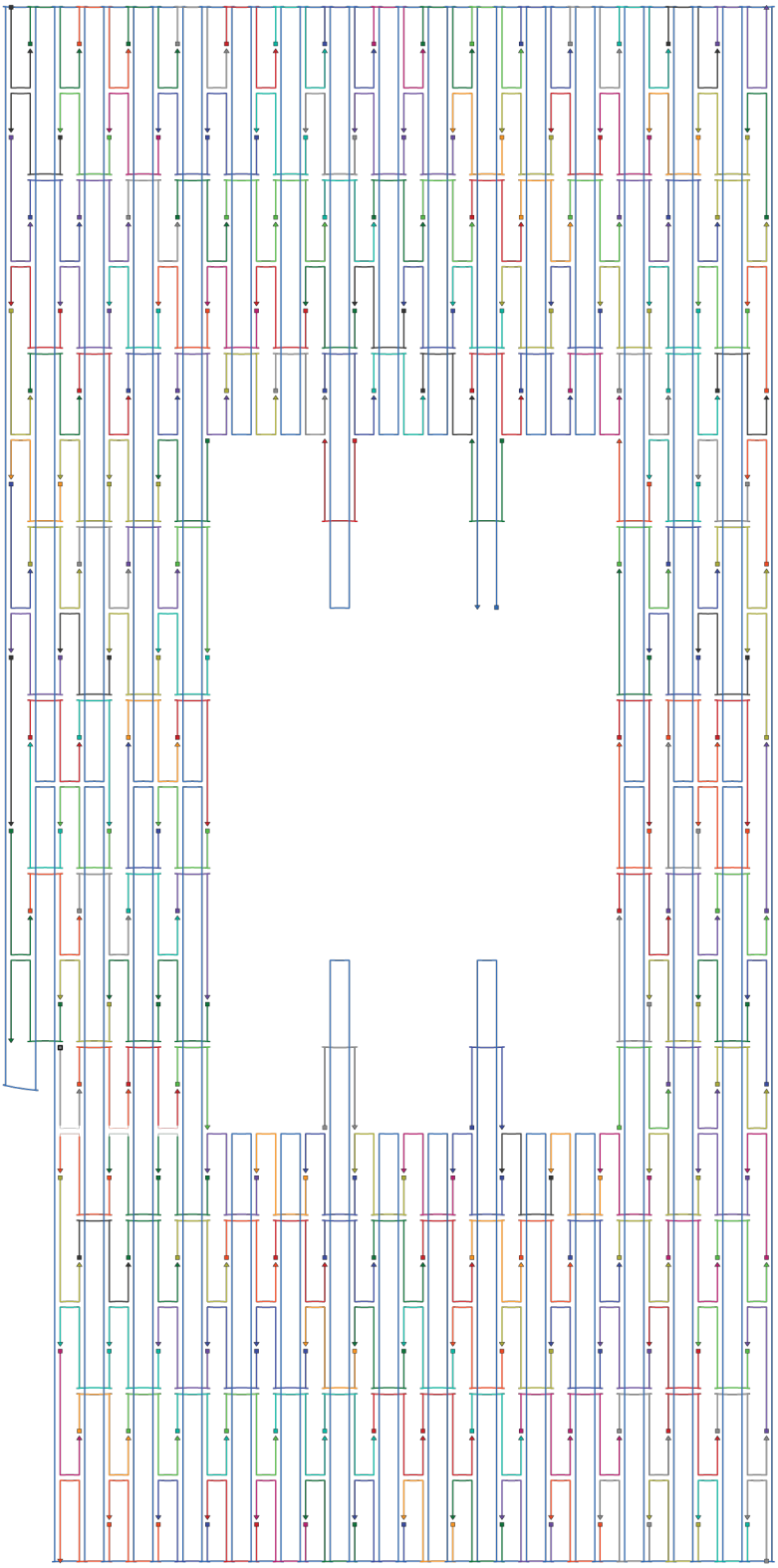

Figure A.1: ***A schematic diagram depicting the scaffold and staple layout of the DF2S origami.***
The M13mp18 scaffold strand is depicted in blue. The direction of each strand is depicted with arrows.

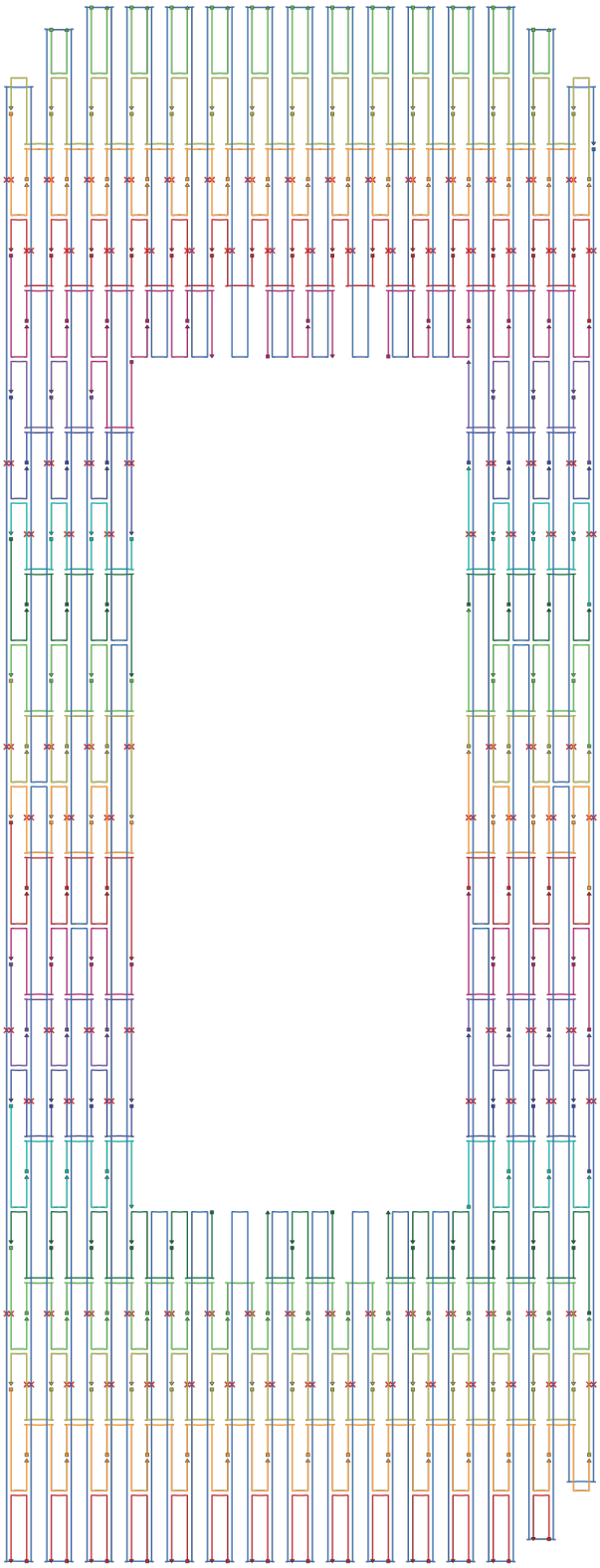

Figure A.2: ***A schematic diagram depicting the scaffold and staple layout of the first version DF2L origami.*** *The M13mp18 scaffold strand is depicted in blue. The direction of each strand is depicted with arrows.*

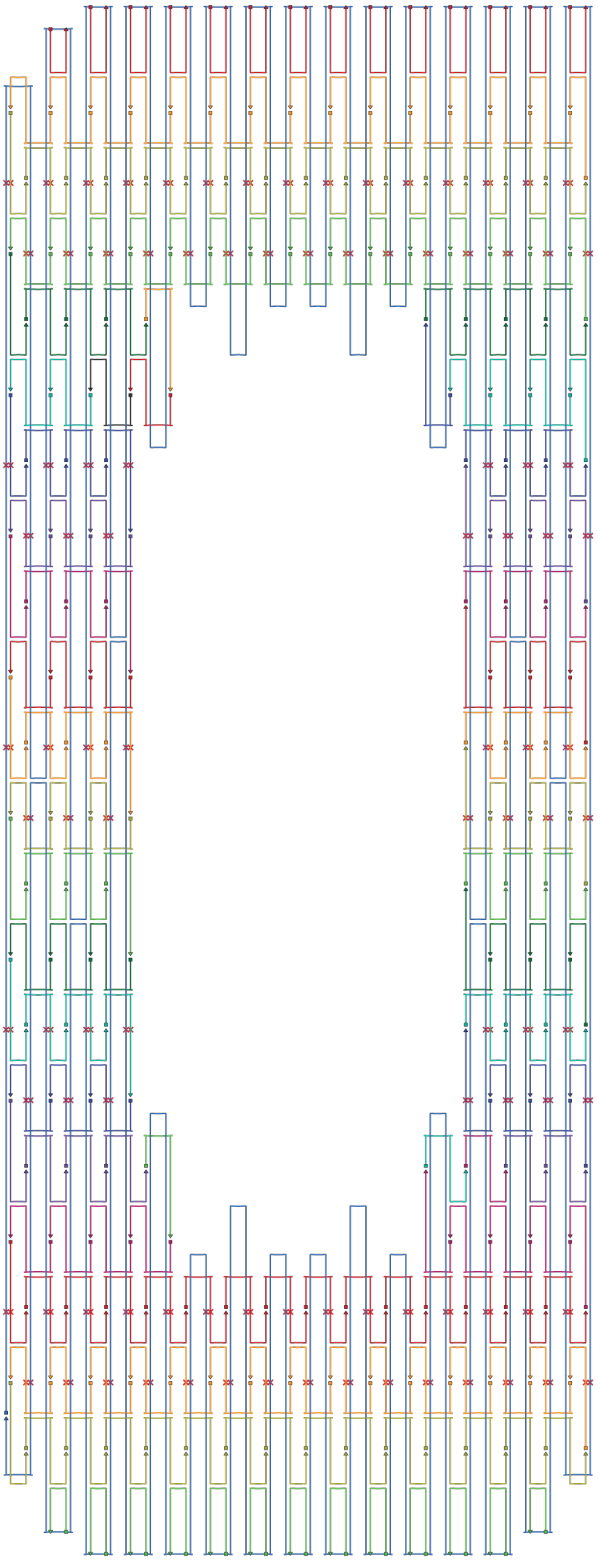

Figure A.3: ***A schematic diagram depicting the scaffold and staple layout of the DF2S origami.*** *The M13mp18 scaffold strand is depicted in blue. The direction of each strand is depicted with arrows.*

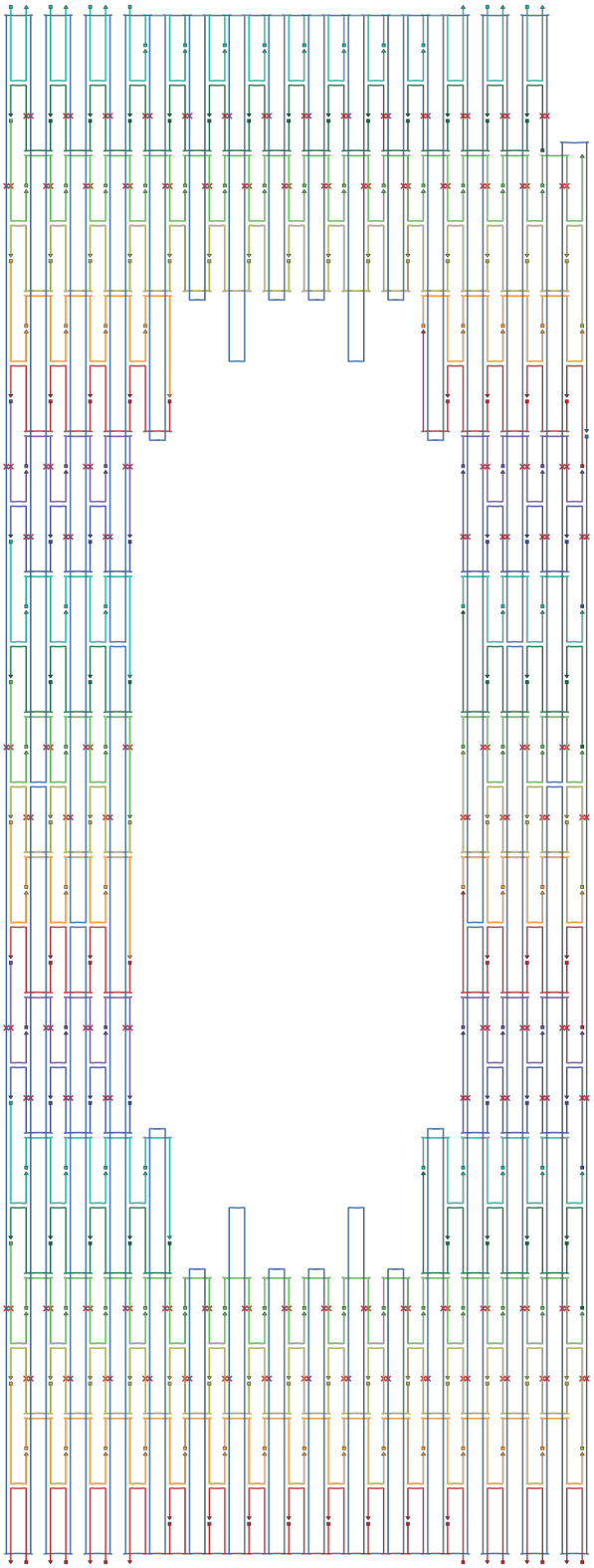

Figure A.4: ***A schematic diagram depicting the scaffold and staple layout of the final version of the DF2L origami.*** *The M13mp18 scaffold strand is depicted in blue. The direction of each strand is depicted with arrows.*

This page is intentionally left blank.

Appendix B

DNA Sequences

In this appendix the sequences of all the oligonucleotides used in this thesis are given. Please note that the oligonucleotides staples that make up the DNA origami frames are excluded due to space limits.

890 bp patterning sequences

890 bp 90 nt patterning

TGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGT

GCCGTAAAGCACTAAATCGGAACCCTAAAGGGAG

890 bp 60 nt patterning

GATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAG

GTGCCGTAAA

890 bp 30 nt patterning

TCAAGTTTTTTGGGGTCGAGGTGCCGTAAA

890 bp Upper primer

GGGGTTCCGCGCACATTTCC

890 bp Lower primer

GGTCGACCTGCAGGCGGCCG

3.5 kbp patterning sequences

3.5 kbp 60 nt patterning

CTGTTCATTATCATCGCTTTTAAAACGGTTCGACCTTCTAATCCTATCTGA

CCATTATAA *3.5 kbp bp Upper primer*

GCCCAAGCCATTAATGGATC

3.5 kbp bp Lower primer

GCAACATGAATAACAGTGGG

DF2S central strand and patterning sequences

Sequence 1: 64 nt Bottom

CTCAAGACGATAGTTACTAGATAAGGAATTCTGGT

CGGGCTGAACGGGGGGTTCGTGCACACAG

Sequence 1: 64 nt Bottom + anchor

GACGGGAGAATTAACTCTCAAGACGATAGTTACTAGATAAGGAA

TTCTGGTCGGGCTGAACGGGGGGTTCGTGCACACAGTTTAAATATGCAACTA

Sequence 1: 64 nt Top

CTGTGTGCACGAACCCCCCGTTCAGCCCGACCA

GAATTCCTTATCTAGTAACTATCGTCTTGAG

Sequence 1: 64 nt Top + anchor

CTGTAGCTCAACATGTCTGTGTGCACGAACCCCCCGTTCA

GCCCGACCAGAATTCCTTATCTAGTAACTATCGTCTTGAGGAACACCCTGAACAAA

Sequence 2: 74 nt Bottom

AGTGAGGAGCAACGCGCACGGATCCATGGTAGGAATTCAAC

AACAATGAATATTTGGAACACTCTAGAGTCTCC

Sequence 2: 74 nt Bottom + anchor

CGACAATAAACAACATAGTGAGGAGCAACGCGCACGGATCCAT

GGTAGGAATTCAACAACAATGAATATTTGGAACACTCTAGAGTCTCCAGCAAACAAGAGAATC

Sequence 2: 74 nt Top

GGAGACTCTAGAGTGTTCCAAATATTCATTGTTGTTGAATTCCTA

CCATGGATCCGTGCGCGTTGCTCCTCACT

Sequence 2: 74 nt Top + anchor

TTGCCTGAGAGTCTGGGGAGACTCTAGAGTGTTCCAAATATTCATT

GTTGTTGAATTCCTACCATGGATCCGTGCGCGTTGCTCCTCACTGTTCAGCTAATGCAGA

Sequence 1: Patterning oligo 30 nt

AACCCCCCGTTCAGCCCGACCAGAATTCCT

Sequence 2: Patterning oligo 30 nt

TTCATTGTTGTTGAATTCCTACCATGGATC

DF2L central strand PCR primers and patterning sequences

Upper sequence A - 224 bp (UA224)

ACGACGATAAAAACCACAGAGACTCATGGAGCAG

Lower sequence A - 224 bp (LA224)

TAAACACCGGAATCATCATGGACTCCCGGG

Upper sequence A - 234 bp (UA234)

ACGACGATAAAAACCACAGAGACTCGCGTGATG

Lower sequence A - 234 bp (LA234)

TAAACACCGGAATCATCATGGACTCTCATACCGG

Upper sequence B - 224 bp (UA224)

TGTTTTAAATATGCAACAGAGACTCCGCAACG

Lower sequence B - 224 bp (LA224)

GATGAAACAAACATCACATGGACTCCTGTGCG

Upper sequence B - 234 bp (UA234)

TGTTTTAAATATGCAACAGAGACTCACAGACGCA

Lower sequence B - 234 bp (LA234)

GATGAAACAAACATCACATGGACTCAGTTCCTGT

Competing oligo - Anchor ALU

ACATAACGCCAAAAGG

Competing oligo - Anchor BRL

GAATAAACACCGGAAT

Competing oligo - Anchor CLU

AGCTCAACATGTTTTA

Competing oligo - Anchor DRL

AAGTTACAAAATCGCG

Sequence A patterning oligo

AAATGCTGAAATGAATTCTAAGCGGAGATC

Sequence B patterning oligo

TGCTTCATCCGCGATATCGCAGTCGGCGTC

Appendix C

Description of movies

In this appendix descriptions of the provided movies are given.

Movie 7.1

A HS-AFM movie depicting the surface mobility of 890 bp dsDNA deposited on Mica that has been pre-incubated with 10 mM $Ni^{2+}$and imagined in Tris buffer containing 10 mM Mg^{2+}.

Movie 7.2

This movie gives a graphical overview of the vector fitting and cross correlation methodology utilised for quantifying the surface translational mobility of DNA molecules.

Movie 7.3

A set of HS-AFM movies depicting the surface mobility of 3.5 kbp dsDNA molecules across a series of Ni^{2+}pre-incubation concentrations, from 15 - 1 mM.

Movie 7.4

A HS-AFM movie depicting the digestion of dsDNA by the restriction enzyme EcoRI. The dsDNA is partially immobilised upon a mica surface pre-incubated with 5 mM $Ni^{2+}$and the reaction conducted in an imaging buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22$^{\circ}C$.

Movie 7.5

A set of HS-AFM movies depicting the formation of RecA nucleation clusters around 3.5 kbp dsDNA in the presence of ATPγS. DNA is immobilised upon a 10 mM Ni^{2+}pre-incubated Mica surface and the reaction conducted in an imaging buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22$^{\circ}C$.

Movie 7.6

A set of HS-AFM movies depicting the surface diffusion of 60 nt RecA nucleoprotein filaments at 22 $^{\circ}C$. Imaging is conducted on a 10 mM Ni^{2+}pre-incubated mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}.

Movie 7.7

Representative HS-AFM movies demonstrating the observed interactions of 60 nt RecA nucleoprotein filaments with 3.5 kbp dsDNA. Imaging is conducted on a 10 mM Ni^{2+}pre-incubated mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 7.8

A HS-AFM depicting the observation of 60 nt RecA nucleoprotein filament clustering behaviour. Imaging is conducted on a 10 mM Ni^{2+}pre-incubated mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 8.1

A series of HS-AFM movies depicting the observed interactions of 30 nt RecA nucleoprotein filaments with dsDNA within a DNA origami frame. Representative examples of frames imaged across a variety orientations with respect to the fastscan axis of the scanning probe are shown. Imaging is conducted on a mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 8.2

A representative HS-AFM movie and derived vector plot depicting the formation of a postsynaptic complex within a DNA origami frame. A 30 nt RecA nucleoprotein filament is observed to remain at the location of homology and move with the dsDNA through successive AFM image. Imaging is conducted on a mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 8.3

A HS-AFM movie and derived vector plot depicting a 30 nt RecA nucleoprotein filament conducting a homology search along non-contiguous sections of dsDNA. Imaging is conducted on a mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 8.4

A HS-AFM movie depicting the interaction of 30 nt RecA nucleoprotein filaments with a portion of a broken DNA origami frame. A filament is observed to conduct a long progressive slide. Imaging is conducted on a mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

Movie 8.5

A series of HS-AFM movies depicting the detrimental flexibility of the enlarged DNA origami frame. Imaging is conducted on a mica surface in a buffer containing 10 mM Tris-AC, pH 7.4 and 10 mM Mg^{2+}at 22 $^{\circ}C$.

This page is intentionally left blank.

Bibliography

1. Seeman, N. C. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, (1982).

2. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, (2006).

3. Gopinath, A. and Rothemund, P. W. K. Optimized Assembly and Covalent Coupling of Single-Molecule DNA Origami Nanoarrays. *ACS Nano*, 8(12):12030–12040, (2014).

4. Sharma, R. Davies, A. G. and Wälti, C. RecA Protein Mediated Nano-Scale Patterning of DNA Scaffolds. *Journal of Nanoscience and Nanotechnology*, 11(12):10629–10632, (2011).

5. Sharma, R. Davies, A. G. and Wälti, C. Nanoscale Programmable Sequence-Specific Patterning of DNA Scaffolds Using RecA Protein. *Nanotechnology*, 23(36):365301, (2012).

6. Sharma, R. Davies, A. G. and Wälti, C. Directed Assembly of 3-nm-Long RecA Nucleoprotein Filaments on Double-Stranded DNA with Nanometer Resolution. *ACS Nano*, 8(4):3322–30, (2014).

7. Cox, M. M. Regulation of Bacterial RecA Protein Function. *Critical Reviews in Biochemistry and Molecular Biology*, 42(1):41–63, (2007).

8. Nishinaka, T. Takano, A. Doi, Y. Hashimoto, M. Nakamura, A. Matsushita, Y. Kumaki, J. and Yashima, E. Conductive Metal Nanowires Templated by the Nucleoprotein Filaments, Complex of DNA and RecA Protein. *Journal of the American Chemical Society*, 127(22):8120–5, (2005).

9. Keren, K. Berman, R. S. Buchstab, E. Sivan, U. and Braun, E. DNA-Templated Carbon Nanotube Field-Effect Transistor. *Science*, 302(5649):1380–2, (2003).

10. Keren, K. Krueger, M. Gilad, R. Ben-Yoseph, G. Sivan, U. and Braun, E. Sequence-Specific Molecular Lithography on Single DNA Molecules. *Science*, 297(5578):72–5, (2002).

11. Forget, A. L. and Kowalczykowski, S. C. Single-Molecule Imaging of DNA Pairing by RecA Reveals a Three-Dimensional Homology Search. *Nature*, 482(7385):423–7, (2012).

12. Müller, B. Koller, T. and Stasiak, A. Characterization of the DNA Binding Activity of Stable RecA-DNA Complexes. Interaction Between the Two DNA Binding Sites Within RecA Helical Filaments. *Journal of Molecular Biology*, 212(1):97–112, (1990).

13. Adzuma, K. No Sliding During Homology Search by RecA Protein. *The Journal of Biological Chemistry*, 273(47):31565–73, (1998).

14. van der Heijden, T. Modesti, M. Hage, S. Kanaar, R. Wyman, C. and Dekker, C. Homologous Recombination in Real Time: DNA Strand Exchange by RecA. *Molecular Cell*, 30(4):530–8, (2008).

15. Ragunathan, K. Liu, C. and Ha, T. RecA Filament Sliding on DNA Facilitates Homology Search. *eLife*, 2012(1):1–14, (2012).

16. Ragunathan, K. Joo, C. and Ha, T. Real-Time Observation of Strand Exchange Reaction with High Spatiotemporal Resolution. *Structure*, 19(8):1064–1073, (2011).

17. Mani, A. Braslavsky, I. Arbel-Goren, R. and Stavans, J. Caught in the Act: The Lifetime of Synaptic Intermediates During the Search for Homology on DNA. *Nucleic Acids Research*, 38(6):2036–2043, (2009).

18. Story, R. M. Weber, I. T. and Steitz, T. A. The Structure of the E. coli RecA Protein Monomer and Polymer. *Nature*, 355(6358):318–25, (1992).

19. Egelman, E. H. and Stasiak, A. Structure of Helical RecA-DNA Complexes Formed in the Presence of ATP-γ-S or ATP. *Journal of Molecular Biology*, 191(4):677–97, (1986).

20. Egelman E. H. and Stasiak, A. Structure of Helical RecA-DNA Complexes. II. Local Conformational Changes Visualized in Bundles of RecA-ATP-γ-S Filaments. *Journal of Molecular Biology*, 200(2):329–49, (1988).

21. Egelman E. H. and Stasiak, A. Electron Microscopy of RecA-DNA Complexes: Two Different States, Their Functional Significance and Relation to the Solved Crystal Structure. *Micron*, 24(3):309–324, (1993).

22. Chen, Z. Yang, H. and Pavletich, N. P. Mechanism of Homologous Recombination from the RecA-ssDNA/dsDNA Structures. *Nature*, 453(7194):489–4, (2008).

23. Alberts, B. Johnson, B. Lewis, J. Raff, M. Roberts, K. and Walter, P. Molecular Biology of the Cell, 5th Edition. *Garland Science Talylor & Francis Group*, (2008).

24. Watson, J. D. and Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356):737–738, (1953).

25. McCarty, M. and Avery, O. T. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: ii. Effect of Desoxyribonuclease on the Biological Activity of the Transforming Substance. *The Journal of Experimental Medicine*, 83(2):89–96, (1946).

26. Rychlik, W. and Rhoads, R. E. A Computer Program for Choosing Optimal Oligonucleotides for Filter Hybridization, Sequencing and In Vitro Amplification of DNA. *Nucleic Acids Research*, 17(21):8543–8551, (1989).

27. Nosenko, Y. Kunitski, M. Stark, T. Göbel, M. Tarakeshwar, P. and Brutschy, B. 4-Aminobenzimidazole-1-methylthymine: a Model for Investigating Hoogsteen Base-Pairing Between Adenine and Thymine. *The Journal of Physical Chemistry. A*, 115(41):11403–11, (2011).

28. Owczarzy, R. You, Y. Groth, C. L. and Tataurov, A. V. Stability and Mismatch Discrimination of Locked Nucleic Acid-DNA Duplexes. *Biochemistry*, 50(43):9352–9367, (2011).

29. Dirks, R. M. Lin, M. Winfree, E. and Pierce, N. A. Paradigms for Computational Nucleic Acid Design. *Nucleic Acids Research*, 32(4):1392–403, (2004).

30. Seeman, N. C. Nanomaterials Based on DNA. *Annual Review of Biochemistry*, 79:65–87, (2010).

31. Neidle, S. *Nucleic Acid Structure and Recognition*. Oxford University Press, USA, (2002).

32. Hsu, H-P. Paul, W. and Binder, K. Breakdown of the Kratky-Porod Wormlike Chain Model for Semiflexible Polymers in Two Dimensions. *EPL (Europhysics Letters)*, 95(6):68004, (2011).

33. Ramachandran, G. N. Ramakrishnan, C. and Sasisekharan, V. Stereochemistry of Polypeptide Chain Configurations. *Journal of Molecular Biology*, 7(1):95–99, (1963).

34. Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science*, 181(96):223–230, (1973).

35. Eisenberg, D. The Discovery of the Alpha-Helix and Beta-Sheet, the Principal Structural Features of Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20):11207–10, (2003).

36. Meiler, J. and Baker, D. Coupled Prediction of Protein Secondary and Tertiary Structure. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):12105–12110, (2003).

37. Bragg, L. and Perutz, M. F. The Structure of Haemoglobin. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 213(1115):425–435, (1952).

38. Seeman, N. C. Structural DNA Nanotechnology: Growing Along with Nano Letters. *Nano letters*, 10(6):1971–1978, (2010).

39. Sa-Ardyen, P. Vologodskii, A. V. and Seeman, N. C. The Flexibility of DNA Double Crossover Molecules. *Biophysical Journal*, 84(6):3829–3837, (2003).

40. Hadden, J. M. Déclais, A-C. Carr, S. B. Lilley, D. M. J. and Phillips, S. E. V. The Structural Basis of Holliday Junction Resolution by T7 Endonuclease I. *Nature*, 449(7162):621–624, (2007).

41. Constantinou, A. Davies, A. A. and West, S. C. Branch Migration and Holliday Junction Resolution Catalyzed by Activities from Mammalian Cells. *Cell*, 104(2):259–268, (2001).

42. Lu, M. Guo, Q. Marky, L. A. and Kallenbach, N. R. Thermodynamic of DNA Branching. *Journal of Molecular Biology*, 223:781–789, (1992).

43. Winfree, E. Liu, F. Wenzler, L. A. and Seeman, N. C. Design and Self-Assembly of Two-Dimensional DNA Crystals. *Nature*, 394(6693):539–544, (1998).

44. Yan, H. Park, S. H. Finkelstein, G. Reif, J. H. and LaBean, T. H. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science*, 301(5641):1882–1884, (2003).

45. He, Y. Tian, Y. Ribbe, A. E. and Mao, C. Highly Connected Two-Dimensional Crystals of DNA Six-Point-Stars. *Journal of the American Chemical Society*, 128(50):15978–15979, (2006).

46. Zheng, J. Birktoft, J. J. Chen, Y. Wang, T. Sha, R. Constantinou, P. E. Ginell, S. L. Mao, C. and Seeman, N. C. From Molecular to Macroscopic via the Rational Design of a Self-Assembled 3D DNA Crystal. *Nature*, 461(7260):74–77, (2009).

47. LaBean, T. H. Yan, H. Kopatsch, J. Liu, F. Winfree, E. Reif, J. H. and Seeman, N. C. Construction, Analysis, Ligation and Self-Assembly of DNA Triple Crossover Complexes. *Journal of the American Chemical Society*, 122(12):1848–1860, (2000).

48. He, Y. Chen, Y. Liu, H. Ribbe, A. E. and Mao, C. Self-Assembly of Hexagonal DNA Two-Dimensional (2D) Arrays. *Journal of the American Chemical Society*, 127(35):12202–12203, (2005).

49. Malo, J. Mitchell, J. C. Vénien-Bryan, C. Harris, J. R. Wille, H. Sherratt, D. J. and Turberfield, A. J. Engineering a 2D Protein-DNA Crystal. *Angewandte Chemie - International Edition*, 44(20):3057–3061, (2005).

50. Majumder, U. Rangnekar, A. Gothelf, K. V. Reif, J. H. and LaBean, T. H. Design and Construction of Double-Decker Tile as a Route to Three-Dimensional Periodic Assembly of DNA. *Journal of the American Chemical Society*, 133(11):3843–3845, (2011).

51. Chen, J. and Seeman, N. C. Synthesis from DNA of a Molecule with the Connectivity of a Cube. *Nature*, 350:631, (1991).

52. Goodman, R. P. Schaap, I. A. T. Tardin, C. F. Erben, C. M. Berry, R. M. Schmidt, C. F. and Turberfield, A. J. Rapid Chiral Assembly of Rigid DNA Building Blocks for Molecular Nanofabrication. *Science*, 310(5754):1661–1665, (2005).

53. He, Y. Ye, T. Su, M. Zhang, C. Ribbe, A. E. Jiang, W. and Mao, C. Hierarchical Self-Assembly of DNA into Symmetric Supramolecular Polyhedra. *Nature*, 452(7184):198–201, (2008).

54. Aldaye, F. A. and Sleiman, H. F. Modular Access to Structurally Switchable 3D Discrete DNA Assemblies. *Journal of the American Chemical Society*, 129(44):13376–13377, (2007).

55. Zhang, Y. W. and Seeman, N. C. Construction of a DNA-Truncated Octahedron. *Journal Of The American Chemical Society*, 116(5):1661–1669, (1994).

56. Shih, W. M. Quispe, J. D. and Joyce, G. F. A 1.7-Kilobase Single-Stranded DNA that Folds into a Nanoscale Octahedron. *Nature*, 427(6975):618–621, (2004).

57. Erben, C. M. Goodman, R. P. and Turberfield, A. J. A Self-Assembled DNA Bi-Pyramid. *Journal of the American Chemical Society*, 129(2):6992–6993, (2007).

58. Yan, H. Park, S. H. Finkelstein, G. Reif, J. H. and LaBean, T. H. DNA-Templated Self-Assembly of Protein Arrays and Highly Conductive Nanowires. *Science*, 301(5641):1882–1884, (2003).

59. Reishus, D. Shaw, B. Brun, Y. Chelyapov, N. and Adleman, L. Self-Assembly of DNA Double-Double Crossover Complexes into High-Density, Doubly Connected, Planar Structures. *Journal of the American Chemical Society*, 127(50):17590–17591, (2005).

60. Zhang, F. Nangreave, J. Liu, Y. and Yan, H. Structural DNA Nanotechnology: State of the Art and Future Perspective. *Journal of the American Chemical Society*, 136(32):11198–11211, (2014).

61. Yan, H. LaBean, T. H. Feng, L. and Reif, J. H. Directed Nucleation Assembly of DNA Tile Complexes for Barcode-Patterned Lattices. *Proceedings of the National Academy of Sciences of the United States of America*, 100(14):8103–8108, (2003).

62. Ding, B. Sha, R. and Seeman, N. C. Pseudo-hexagonal 2D DNA Crystals from Double Crossover Cohesion. *Journal of the American Chemical Society*, 126(33):10230–10231, (2004).

63. Constantinou, P. E. Wang, T. Kopatsch, J. Israel, L. B. Zhang, X. Ding, B. Sherman, W. B. Wang, X. Zheng, J. Sha, R. and Seeman, N. C. Double Cohesion in Structural DNA Nanotechnology. *Organic & Biomolecular Chemistry*, 4(18):3414–3419, (2006).

64. Rothemund, P. W. K. Design of DNA Origami. *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, 2005:470–477, (2005).

65. Castro, C. E. Kilchherr, F. Kim, D-N. Shiao, E. L. Wauer, T. Wortmann, P. Bathe, M. and Dietz, H. A Primer to Scaffolded DNA Origami. *Nature Methods*, 8(3):221–229, (2011).

66. Douglas, S. M. Dietz, H. Liedl, T. Högberg, B. Graf, F. and Shih, W. M. Self-Assembly of DNA into Nanoscale Three-Dimensional Shapes. *Nature*, 459(7245):414–418, (2009).

67. Ke, Y. Sharma, J. Liu, M. Jahn, K. Liu, Y. and Yan, H. Scaffolded DNA Origami of a DNA Tetrahedron Molecular Container. *Nano Letters*, 9(6):2445–2447, (2009).

68. Andersen, E. S. Dong, M. Nielsen, M. M. Jahn, K. Subramani, R. Mamdouh, W. Golas, M. M. Sander, B. Stark, H. Oliveira, C. L. P. Pedersen, J. S. Birkedal, V. Besenbacher, F. Gothelf, K. V. and Kjems, J. Self-Assembly of a Nanoscale DNA Box with a Controllable lid. *Nature*, 459(7243):73–76, (2009).

69. Endo, M. Hidaka, K. Kato, T. Namba, K. and Sugiyama, H. DNA Prism Structures Constructed by Folding of Multiple Rectangular Arms. *Journal of the American Chemical Society*, 131(43):15570–15571, (2009).

70. Zhao, Z. Liu, Y. and Yan, H. Organizing DNA Origami Tiles into Larger Structures Using Preformed Scaffold Frames. *Nano Letters*, 11(7):2997–3002, (2011).

71. Woo, S. and Rothemund, P. W. K. Self-Assembly of Two-Dimensional DNA Origami Lattices Using Cation-Controlled Surface Diffusion. *Nature Communications*, 5:4889, 2014.

72. Marchi, A. N. Saaem, I. Vogen, B. N. Brown, S. and LaBean, T. H. Toward Larger DNA Origami. *Nano Letters*, 14(10):5740–5747, (2014).

73. Ke, Y. Douglas, S. M. Liu, M. Sharma, J. Cheng, A. Leung, A. Liu, Y. Shih, W. M. and Yan. H. Multilayer DNA Origami Packed on a Square Lattice. *Journal of the American Chemical Society*, 131(43):15903–15908, (2009).

74. Dietz, H. Douglas, S. M. and Shih, W. M. Folding DNA into Twisted and Curved Nanoscale Shapes. *Science*, 325(5941):725–730, (2009).

75. Douglas, S. M. Marblestone, A. H. Teerapittayanon, S. Vazquez, A. Church, G. M. and Shih, W. M. Rapid Prototyping of 3D DNA-Origami Shapes with caDNAno. *Nucleic Acids Research*, 37(15):5001–5006, (2009).

76. Geary, C. Rothemund, P. W. K. and Andersen, E. S. A Single-Stranded Architecture for Co-Transcriptional Folding of RNA Nanostructures. *Science*, 345(6198):799–804, (2014).

77. Yin, P. Hariadi, R. F. Sahu, S. Choi, H. M. T. Park, S. H. LaBean, T. H. and Reif, J. H. Programming DNA Tube Circumferences. *Science*, 321(5890):824–826, (2008).

78. Wei, B. Dai, M. and Yin, P. Complex Shapes Self-Assembled from Single-Stranded DNA Tiles. *Nature*, 485(7400):623–626, (2012).

79. Ke, Y. Ong, L. L. Shih, W. M. and Yin, P. Three-Dimensional Structures Self-Assembled from DNA bricks. *Science*, 338(6111):1177–1183, (2012).

80. Benson, E. Mohammed, A. Gardell, J. Masich, S. Czeizler, E. Orponen, P. and Högberg, B. DNA Rendering of Polyhedral Meshes at the Nanoscale. *Nature*, 523(7561):441–444, (2015).

81. Zhang, F. Jiang, S. Wu, S. Li, Y. Mao, C. Liu, Y. and Yan, H. Complex Wireframe DNA Origami Nanostructures with Multi-Arm Junction Vertices. *Nature Nanotechnology*, 10(9):779–784, (2015).

82. Han, D. Pal, S. Yang, Y. Jiang, S. Nangreave, J. Liu, Y. and Yan, H. DNA Gridiron Nanostructures Based on Four-Arm Junctions. *Science*, 339(6126):1412–1415, (2013).

83. Andersen, E. S. Biomolecular Self-Assembly: DNA Origami Rewired. *Nature Nanotechnology*, 10(9):733–734, (2015).

84. Tan, S. J. Campolongo, M. J. Luo, D. and Cheng, W. Building Plasmonic Nanostructures with DNA. *Nature Nanotechnology*, 6(5):268–276, (2011).

85. Klein, W. P. Schmidt, C. N. Rapp, B. Takabayashi, S. Knowlton, W. B. Lee, J. Yurke, B. Hughes, W. L. Graugnard, E. and Kuang, W. Multiscaffold DNA Origami Nanoparticle Waveguides. *Nano Letters*, 13(8):3850–3856, (2013).

86. Ding, B. Deng, Z. Yan, H. Cabrini, S. Zuckermann, R. N. and Bokor, J. Gold Nanoparticle Self-Similar Chain Structure Organized by DNA Origami. *Journal of the American Chemical Society*, 132:3248–3249, (2010).

87. Kuzyk, A. Schreiber, R. Fan, Z. Pardatscher, G. Roller, E-M. Högele, A. Simmel, F. C. Govorov, A. O. and Liedl, T. DNA-Based Self-Assembly of Chiral Plasmonic Nanostructures with Tailored Optical Response. *Nature*, 483(7389):311–314, (2012).

88. Pal, S. Deng, Z. Wang, H. Zou, S. Liu, Y. and Yan, H. DNA Directed Self-Assembly of Anisotropic Plasmonic Nanostructures. *Journal of the American Chemical Society*, 133(44):17606–17609, (2011).

89. Deng, Z. Samanta, A. Nangreave, J. Yan, H. and Liu, Y. Robust DNA-Functionalized Core/Shell Quantum Dots with Fluorescent Emission Spanning from UV-Vis to Near-IR and Compatible with DNA-Directed Self-Assembly. *Journal of the American Chemical Society*, 134(42):17424–17427, (2012).

90. Zheng, J. Constantinou, P. E. Micheel, C. Alivisatos, A. P. Kiehl, R. A. and Seeman, N. C. Two-Dimensional Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs. *Nano Letters*, 6(7):1502–1504, (2006).

91. Dutta, P. K. Varghese, R. Nangreave, J. Lin, S. Yan, H. and Liu, Y. DNA-Directed Artificial Light-Harvesting Antenna. *Journal of the American Chemical Society*, 133(31):11985–11993, (2011).

92. Acuna, G. P. Möller, F. M. Holzmeister, P. Beater, S. Lalkens, B. and Tinnefeld, P. Fluorescence Enhancement at Docking Sites of DNA-Directed Self-Assembled Nanoantennas. *Science*, 338(6106):506–510, (2012).

93. Pal, S. Dutta, P. Wang, H. Deng, Z. Zou, S. Yan, H. and Liu, Y. Quantum Efficiency Modification of Organic Fluorophores Using Gold Nanoparticles on DNA Origami Scaffolds. *Journal of Physical Chemistry C*, 117(24):12735–12744, (2013).

94. Chhabra, R. Sharma, J. Wang, H. Zou, S. Lin, S. Yan, H. Lindsay, S. and Liu, Y. Distance-Dependent Interactions Between Gold Nanoparticles and Fluorescent Molecules with DNA as Tunable Spacers. *Nanotechnology*, 20(48):485201, (2009).

95. Jin, Z. Sun, W. Ke, Y. Shih, C-J. Paulus, G. L. C. Wang, Q. H. Mu, B. Yin, P. and Strano, M. S. Metallized DNA Nanolithography for Encoding and Transferring Spatial Information for Graphene Patterning. *Nature Communications*, 4:1663, (2013).

96. Surwade, S. P. Zhao, S. and Liu, H. Molecular Lithography Through DNA-Mediated Etching and Masking of SiO_2. *Journal of the American Chemical Society*, 133(31):11868–11871, (2011).

97. Surwade, S. P. Zhou, F. Wei, B. Sun, W. Powell, A. O'Donnell, C. Yin, P. and Liu, H. Nanoscale Growth and Patterning of Inorganic Oxides Using DNA Nanostructure Templates. *Journal of the American Chemical Society*, 135:6778–6781, (2013).

98. Helmi, S. Ziegler, C. Kauert, D. J. and Seidel, R. Shape-Controlled Synthesis of Gold Nanostructures Using DNA Origami Moulds. *Nano Letters*, 14(11):6693–6698, (2014).

99. Pearson, A. C. Pound, E. Woolley, A. T. Linford, M. R. Harb, J. N. and Davis, R. C. Chemical Alignment of DNA Origami to Block Co-Polymer Patterned Arrays of 5 nm Gold Nanoparticles. *Nano Letters*, 11(5):1981–1987, (2011).

100. Kershner, R. J. Bozano, L. D. Micheel, C. M. Hung, A. M. Fornof, A. R. Cha, J. N. Rettner, C. T. Bersani, M. Frommer, J. Rothemund, P. W. K. and Wallraff, G. M. Placement and Orientation of Individual DNA Shapes on Lithographically Patterned Surfaces. *Nature Nanotechnology*, 4(9):557–561, (2009).

101. Aren, E. G. Soo, O. S. Kuangwen, H. Yonggang, K. Hao, Y. and Soh, H. T. Controlled Delivery of DNA Origami on Patterned Surfaces. *Small*, 5(17):1942–1946, (2009).

102. Hung, A. M. Micheel, C. M. Bozano, L. D. Osterbur, L. W. Wallraff, G. M. and Cha, J. N. Large-Area Spatially Ordered Arrays of Gold Nanoparticles Directed by Lithographically Confined DNA Origami. *Nature Nanotechnology*, 5(2):121–126, (2010).

103. Hung, A. M. and Cha, J. N. Templated Assembly of DNA Origami Gold Nanoparticle Arrays on Lithographically Patterned Surfaces. *Methods in Molecular Biology*, 749:187–197, (2011).

104. Teshome, B. Facsko, S. and Keller, A. Topography-Controlled Alignment of DNA Origami Nanotubes on Nanopatterned Surfaces. *Nanoscale*, 6(3):1790–1796, (2014).

105. Chhabra, R. Sharma, J. Ke, Y. Liu, Y. Rinker, S. Lindsay, S. and Yan, H. Spatially Addressable Multiprotein Nanoarrays Templated by Aptamer-Tagged DNA Nanoarchitectures. *Journal of the American Chemical Society*, 129(34):10304–10305, (2007).

106. Selmi, D. N. Adamson, R. J. Attrill, H. Goddard, A. D. Gilbert, R. J. C. Watts, A. and Turberfield, A. J. DNA-templated Protein Arrays for Single-Molecule Imaging. *Nano Letters*, 11(2):657–660, (2011).

107. Berardi, M. J. Shih, W. M. Harrison, S. C. and Chou, J. J. Mitochondrial Uncoupling Protein 2 Structure Determined by NMR Molecular Fragment Searching. *Nature*, 476(7358):109–113, (2011).

108. Delebecque, C. J. Lindner, A. B. Silver, P. A. and Aldaye, F. A. Organization of Intracellular Reactions with Rationally Designed RNA Assemblies. *Science*, 333(6041):470–474, (2011).

109. Erkelenz, M. Kuo, C. H. and Niemeyer, C. M. DNA-Mediated Assembly of Cytochrome P450 BM3 Subdomains. *Journal of the American Chemical Society*, 133(40):16111–16118, (2011).

110. Wilner, O. I. Weizmann, Y. Gill, R. Lioubashevski, O. Freeman, R. and Willner, I. Enzyme Cascades Activated on Topologically Programmed DNA Scaffolds. *Nature Nanotechnology*, 4(4):249–254, (2009).

111. Fu, J. Liu, M. Liu, Y. Woodbury, N. W. and Yan, H. Inter-Enzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *Journal of the American Chemical Society*, 134(12):5516–5519, (2012).

112. Fu, J. Yang, Y. R. Johnson-Buck, A. Liu, M. Liu, Y. Walter, N. G. Woodbury, N. W. and Yan, H. Multi-Enzyme Complexes on DNA Scaffolds Capable of Substrate Channelling with an Artificial Swinging Arm. *Nature Nanotechnology*, 9(7):531–536, (2014).

113. Derr, N. D. Goodman, B. S. Jungmann, R. Leschziner, A. E. Shih, W. M. and Reck-Peterson, S. L. Tug-of-War in Motor Protein Ensembles. *Science*, 338(2):662–666, (2012).

114. Qiu, W. Derr, N. D. Goodman, B. S. Villa, E. Wu, D. Shih, W. M. and Reck-Peterson, S. L. Dynein Achieves Processive Motion Using Both Stochastic and Coordinated Stepping. *Nature Structural & Molecular Biology*, 19(2):193–200, (2012).

115. Hariadi, R. F. Cale, M. and Sivaramakrishnan, S. Myosin Lever Arm Directs Collective Motion on Cellular Actin Network. *Proceedings of the National Academy of Sciences of the United States of America*, 111(11):4091–4096, (2014).

116. Pfitzner, E. Wachauf, C. Kilchherr, F. Pelz, B. Shih, W. M. Rief, M. and Dietz, H. Rigid DNA Beams for High-Resolution Single-Molecule Mechanics. *Angewandte Chemie - International Edition*, 52(30):7766–7771, (2013).

117. Prinz, J. Schreiber, B. Olejko, L. Oertel, J. Rackwitz, J. Keller, A. and Bald, I. DNA Origami Substrates for Highly Sensitive Surface-Enhanced Raman Scattering. *The Journal of Physical Chemistry Letters*, 4(23):4140–4145, (2013).

118. Thacker, V. V. Herrmann, L. O. Sigle, D. O. Zhang, T. Liedl, T. Baumberg, J. J. and Keyser, U. F. DNA Origami Based Assembly of Gold Nanoparticle Dimers for Surface-Enhanced Raman Scattering. *Nature Communications*, 5:3448, (2014).

119. Voigt, N. V. Tørring, T. Rotaru, A. Jacobsen, M. F. Ravnsbaek, J. B. Subramani, R. Mamdouh, W. Kjems, J. Mokhir, A. Besenbacher, F. and Gothelf, K. V. Single-Molecule Chemical Reactions on DNA Origami. *Nature Nanotechnology*, 5(3):200–203, (2010).

120. Suzuki, Y. Endo, M. Cañas, C. Ayora, S. Alonso, J. C. Sugiyama, H. and Takeyasu, K. Direct Analysis of Holliday Junction Resolving Enzyme in a DNA Origami Nanostructure. *Nucleic Acids Research*, 42(11):7421–7428, (2014).

121. Sannohe, Y. Endo, M. Katsuda, Y. Hidaka, K. and Sugiyama, H. Visualization of Dynamic Conformational Switching of the G-quadruplex in a DNA Nanostructure. *Journal of the American Chemical Society*, 132(46):16311–16313, (2010).

122. Yamamoto, S. De, D. Hidaka, K. Kim, K. K. Endo, M. and Sugiyama, H. Single Molecule Visualization and Characterization of Sox2-Pax6 Complex Formation on a Regulatory DNA Element Using a DNA Origami Frame. *Nano Letters*, 14(5):2286–2292, (2014).

123. Suzuki, Y. Endo, M. and Sugiyama, H. Studying RNAPPromoter Interactions Using Atomic Force Microscopy. *Methods*, 86:4–9, (2015).

124. Endo, M. Takeuchi, Y. Suzuki, Y. Emura, T. Hidaka, K. Wang, F. Willner, I. and Sugiyama, H. Single-Molecule Visualization of the Activity of a Zn¡sup¿2+¡/sup¿-Dependent DNAzyme. *Angewandte Chemie – International Edition*, 54(36):10550–10554, (2015).

125. Rajendran, A. Endo, M. Hidaka, K. Teulade-Fichou, M-P. Mergny, J-L. and Sugiyama, H. Small Molecule Binding to G-Hairpin and G-Triplex: A New Insight in Anti-Cancer Drug Design Targeting G-Rich Regions. *Chemical Communication*, 51:9181–9184, (2015).

126. Endo, M. Inoue, M. Suzuki, Y. Masui, C. Morinaga, H. Hidaka, K. and Sugiyama, H. Regulation of B-Z Conformational Transition and Complex Formation with a Z-form Binding Protein by Introduction of Constraint to Double-Stranded DNA by Using a DNA Nanoscaffold. *Chemistry*, 19(50):16887–16890, (2013).

127. Endo, M. Tatsumi, K. Terushima, K. Katsuda, Y. Hidaka, K. Harada, Y. and Sugiyama, H. Direct Visualization of the Movement of a Single T7 RNA Polymerase and Transcription on a DNA Nanostructure. *Angewandte Chemie – International Edition*, 51(35):8778–8782, (2012).

128. Endo, M. Katsuda, Y. Hidaka, K. and Sugiyama, H. A Versatile DNA Nanochip for Direct Analysis of DNA Base-Excision Repair. *Angewandte Chemie – International Edition*, 49(49):9412–9416, (2010).

129. Linko, V. Ora, A. and Kostiainen, M. A. DNA Nanostructures as Smart Drug-Delivery Vehicles and Molecular Devices. *Trends in Biotechnology*, 33(10):586–594, (2015).

130. Zhao, Y-X. Shaw, A. Zeng, X. Benson, E. and Nystro, A. M. DNA Origami Delivery System for Cancer Therapy with Tunable Release Properties. *ACS Nano*, 6(10):8684–8691, (2012).

131. Jiang, Q. Song, C. Nangreave, J. Liu, X. Lin, L. Qiu, D. Wang, Z. G. Zou, G. Liang, X. Yan, H. and Ding, B. DNA Origami as a Carrier for Circumvention of Drug Resistance. *Journal of the American Chemical Society*, 134(32):13396–13403, (2012).

132. Zhang, Q. Jiang, Li, N. Dai, L. Liu, Q. Song, L. Wang, J. Li, Y. Tian, J. Ding, B. and Du, Y. DNA Origami as an In Vivo Drug Delivery Vehicle for Cancer Therapy. *ACS Nano*, 8(7):6633–6643, (2014).

133. Kumar, V. Bayda, S. Hadla, M. Caligiuri, I. Spena, C. R. Palazzolo, S. Kempter, S. Corona, G. Toffoli, G. and Rizzolio, F. Enhanced Chemotherapeutic Behaviour of Open-Caged DNA@Doxorubicin Nanostructures for Cancer Cells. *Journal of Cellular Physiology*, 231(1):106–110, (2016).

134. Kocabey, S. Meinl, H. MacPherson, I. Cassinelli, V. Manetto, A. Rothenfusser, S. Liedl, T. and Lichtenegger, F. Cellular Uptake of Tile-Assembled DNA Nanotubes. *Nanomaterials*, 5(1):47–60, (2014).

135. Lee, H. Lytton-Jean, A. K. R. Chen, Y. Love, K. T. Park, A. I. Karagiannis, E. D. Sehgal, A. Querbes, W. Zurenko, C. S. Jayaraman, M. Peng, C. G. Charisse, K. Borodovsky, A. Manoharan, M. Donahoe, J. S. Truelove, J. Nahrendorf, M. Langer, R. and Anderson, D. G. Molecularly Self-Assembled Nucleic Acid Nanoparticles for Targeted in vivo siRNA Delivery. *Nature Nanotechnology*, 7(6):389–393, (2012).

136. Schüller, V. J. Heidegger, S. Sandholzer, N. Nickels, P. C. Suhartha, N. A. Endres, S. Bourquin, C. and Liedl, T. Cellular Immunostimulation by CpG-Sequence-Coated DNA Origami Structures. *ACS Nano*, 5(12):9696–9702, (2011).

137. Li, J. Pei, H. Zhu, B. Liang, L. Wei, M. He, Y. Chen, N. Li, D. Huang, Q. and Fan, C. Self-Assembled Multivalent DNA Nanostructures for Non-Invasive Intracellular Delivery of Immunostimulatory CpG Oligonucleotides. *ACS Nano*, 5(11):8783–8789, (2011).

138. Uno, S. Nishikawa, M. Mohri, K. Umeki, Y. Matsuzaki, N. Takahashi, Y. Fujita, H. Kadowaki, N. and Takakura, Y. Efficient Delivery of Immuno-Stimulatory DNA to Mouse and Human Immune Cells Through the Construction of Polypod-like Structured DNA. *Nanomedicine: Nanotechnology, Biology, and Medicine*, 10(4):765–774, (2014).

139. Douglas, S. M. Bachelet, I. and Church, G. M. A Logic-Gated Nanorobot for Targeted Transport of Molecular Payloads. *Science*, 335(6070):831–834, (2012).

140. Kuzuya, A. and Komiyama, M. Design and Construction of a Box-Shaped 3D-DNA Origami. *Chemical Communications*, 28:4182–4184, (2009).

141. West, S. C. Cassuto, E. and Howard-Flanders, P. RecA Protein Promotes Homologous-Pairing and Strand-Exchange Reactions Between Duplex DNA Molecules. *Proceedings of the National Academy of Sciences*, 78(4):2100–2104, (1981).

142. Seitz, E. M. Haseltine, C. A. and Kowalczykowski, S. C. *Archaea: Ancient Microbes, Extreme Environments, and the Origin of Life*, volume 50 of *Advances in Applied Microbiology*. Elsevier, (2001).

143. Khanna, K. K. and Jackson, S. P. DNA Double-Strand Breaks: Signaling, Repair and the Cancer Connection. *Nature Genetics*, 27(3):247–254, (2001).

144. Krejci, L. Altmannova, V. Spirek, M. and Zhao, X. Homologous Recombination and its Regulation. *Nucleic Acids Research*, 40(13):5795–5818, (2012).

145. Hong, S. Sung, Y. Yu, M. Lee, M. Kleckner, N. and Kim, K. P. The Logic and Mechanism of Homologous Recombination Partner Choice. *Molecular Cell*, 51(4):440–453, (2013).

146. Kawabata, M. Kawabata, T. and Nishibori, M. Role of RecA / RAD51 Family of Proteins in Mammals. *Acta Medica Okayama*, 59(1):1–9, (2005).

147. Stasiak, A. Z. Rosselli, W. and Stasiak, A. RecA-DNA Helical Filaments in Genetic Recombination. *Biochimie*, 73:199–208, (1991).

148. Simonson, T. Kubista, M. Sjöback, R. Ryberg, H. and Takahashi, M. Properties of RecA-Oligonucleotide Complexes. *Journal of Molecular Recognition*, 7(3):199–206, (1994).

149. Datta, S. Ganesh, N. Chandra, N. R. Muniyappa, K. and Vijayan, M. Structural Studies on MtRecA-Nucleotide Complexes: Insights into DNA and Nucleotide Binding and the Structural Signature of NTP Recognition. *Proteins*, 50(3):474–485, (2003).

150. Benedict, R. C. and Kowalczykowski, S. C. Increase of the DNA Strand Assimilation Activity of RecA Protein by Removal of the C Terminus and Structure-Function Studies of the Resulting Protein Fragment. *The Journal of Biological Chemistry*, 263(30):15513–15520, (1988).

151. Story, R. M. and Steitz, T. A. Structure of the RecA Protein-ADP Complex. *Nature*, 356:133–135, (1992).

152. van Loenhout, M. T. J. van der Heijden, T. Kanaar, R. Wyman, C. and Dekker, C. Dynamics of RecA Filaments on Single-Stranded DNA. *Nucleic Acids Research*, 37(12):4089–4099, (2009).

153. Nastri, H. G. and Knight, K. L. Identification of Residues in the L1 Region of the RecA Protein Which are Important to Recombination or Coprotease Activities. *The Journal of Biological Chemistry*, 269(42):26311–26322, (1994).

154. Hortnagel, K. Voloshin, O. N. Kinal, H. H. Ma, N. Schaffer-Judge, C. and Camerini-Otero, R. D. Saturation Mutagenesis of the E. coli RecA Loop L2 Homologous DNA Pairing Region Reveals Residues Essential for Recombination and Recombinational Repair. *Journal of Molecular Biology*, 286(4):1097–1106, (1999).

155. Skibas, M. C. and Knight, K. L. Functionally Important Residues at a Subunit Interface Site. *Journal of Biological Chemistry*, 269(5):3823–3828, (1994).

156. Malkov, V. A. and Camerini-Otero, R. D. Photocross-Links Between Single-Stranded DNA and Escherichia coli RecA Protein Map to Loops L1 (Amino Acid Residues 157-164) and L2 (Amino Acid Residues 195-209). *The Journal of Biological Chemistry*, 270(50):30230–30233, (1995).

157. Voloshin, O. N. Wang, L. and Camerini-Otero, R. D. Homologous DNA Pairing Promoted by a 20-Amino Acid Peptide Derived from RecA. *Science*, 272(5263):868–872, (1996).

158. Kumars, K. A. Mahalakshmis, S. and Muniyappasoll, K. DNA-Induced Conformational Changes in RecA Protein. *Journal of Biological Chemistry*, 268(35):26162–26170, (1993).

159. Chen, Z. Yang, H. and Pavletich, N. P. Mechanism of Homologous Recombination from the RecA-ssDNA/dsDNA Structures. *Nature*, 453(7194):489–494, (2008).

160. Yang, D. Boyer, B. Prévost, C. Danilowicz, C. and Prentiss, M. Integrating Multi-Scale Data on Homologous Recombination into a New Recognition Mechanism Based on Simulations of the RecA-ssDNA/dsDNA Structure. *Nucleic Acids Research*, 43(21):10251–10263, (2015).

161. Saladin, A. Amourda, C. Poulain, P. Férey, N. Baaden, M. Zacharias, M. Delalande, O. and Prévost, C. Modelling the Early Stage of DNA Sequence Recognition Within RecA Nucleoprotein Filaments. *Nucleic Acids Research*, 38(19):6313–6323, (2010).

162. Konola, J. T. Logan, K. M. and Knight, K. L. Functional Characterization of Residues in the P-loop Motif of the RecA Protein ATP Binding Site. *Journal of Molecular Biology*, 237(1):20–34, (1994).

163. Kowalczykowski, S. C. Dixon, D. A. Eggleston, A. K. Lauder, S. D. and Rehrauer, W. M. Biochemistry of Homologous Recombination in Escherichia coli. *Microbiological Reviews*, 58(3):401–465, (1994).

164. Lusetti, S. L. and Cox, M. M. The Bacterial RecA Protein and the Recombinational DNA Repair of Stalled Replication Forks. *Annual Review of Biochemistry*, 71:71–100, (2002).

165. Rehrauer, W. M. and Kowalczykowski, S. C. Alteration of the Nucleoside Triphosphate (NTP) Catalytic Domain Within Escherichia coli RecA Protein Attenuates NTP Hydrolysis but not Joint Molecule Formation. *Journal of Biological Chemistry*, 268(2):1292–1297, (1993).

166. Shan, Q. Cox, M. M. and Inman, R. B. DNA Strand Exchange Promoted by RecA K72R. *Journal of Biological Chemistry*, 271(10):5712–5724, (1996).

167. Campbell, M. J. and Davis, R. W. On the in vivo Function of the RecA ATPase. *Journal of Molecular Biology*, 286(2):437–445, (1999).

168. Yu, X. Jacobs, S. A. West, S. C. Ogawa, T. and Egelman, E. H. Domain Structure and Dynamics in the Helical Filaments Formed by RecA and Rad51 on DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 98(15):8419–8424, (2001).

169. Tateishi, S. Horii, T. Ogawa, T. and Ogawa, H. C-terminal Truncated Escherichia coli RecA Protein RecA5327 has Enhanced Binding Affinities to Single- and Double-Stranded DNAs. *Journal of Molecular Biology*, 223(1):115–129, (1992).

170. Knight, K. L. Aoki, K. H. Ujita, E. L. and McEntee, K. Identification of the Amino Acid Substitutions in Two Mutant Forms of the RecA Protein From Escherichia coli: RecA441 and RecA629. *The Journal of Biological Chemistry*, 259(18):11279–11283, (1984).

171. Lavery, P. E. and Kowalczykowski, S. C. Biochemical Basis of the Temperature-Inducible Constitutive Protease Activity of the RecA441 Protein of Escherichia coli. *Journal of Molecular Biology*, 203(4):861–874, (1988).

172. Brenner, S. L. Mitchell, R. S. Morrical, S. W. Neuendorf, S. K. Schutte, B. C. and Cox, M. M. RecA Protein-Promoted ATP Hydrolysis Occurs Throughout RecA Nucleoprotein Filaments. *The Journal of Biological Chemistry*, 262(9):4011–4016, (1987).

173. Savir, Y. and Tlusty, T. RecA-Mediated Homology Search as a Nearly Optimal Signal Detection System. *Molecular Cell*, 40(3):388–396, (2010).

174. Yu, X. and Egelman, E. H. The RecA Hexamer is a Structural Homologue of Ring Helicases. *Nature Structural Biology*, 4(2):101–104, (1997).

175. Mikawa, T. Masui, R. Ogawa, T. Ogawa, H. and Kuramitsu, S. N-terminal 33 Amino Acid Residues of Escherichia coli RecA Protein Contribute to its Self-Assembly. *Journal of Molecular Biology*, 250(4):471–483, (1995).

176. Revet, B. M. J. Sena, E. P. and Zarling, D. A. Homologous DNA Targeting with RecA Protein-Coated Short DNA Probes and Electron Microscope Mapping on Linear Duplex Molecules. *Journal of Molecular Biology*, 232(3):779–791, (1993).

177. Renkawitz, J. Lademann, C. A. and Jentsch, S. Mechanisms and Principles of Homology Search During Recombination. *Nature Reviews. Molecular Cell Biology*, 15(6):369–383, (2014).

178. Galletto, R. Amitani, I. Baskin, R. J. and Kowalczykowski, S. C. Direct Observation of Individual RecA Filaments Assembling on Single DNA Molecules. *Nature*, 443(7113):875–878, (2006).

179. Kowalczykowski, S. C. Clow, J. and Krupp, R. A. Properties of the Duplex DNA-Dependent ATPase Activity of Escherichia coli RecA Protein and its Role in Branch Migration. *Proceedings of the National Academy of Sciences of the United States of America*, 84:3127–3131, (1987).

180. Bar-Ziv, R. and Libchaber, A. Effects of DNA Sequence and Structure on Binding of RecA to Single-Stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 98(16):9068–9073, (2001).

181. Sugiyama, T. Zaitseva, E. M. and Kowalczykowski, S. C. A Single-Stranded DNA-Binding Protein is Needed for Efficient Presynaptic Complex Formation by the Saccharomyces cerevisiae Rad51 Protein. *Journal of Biological Chemistry*, 272(12):7940–7945, (1997).

182. Pugh, B. F. and Cox, M. M. General Mechanism for RecA Protein Binding to Duplex DNA. *Journal of Molecular Biology*, 203(2):479–493, (1988).

183. Joo, C. McKinney, S. A. Nakamura, M. Rasnik, I. Myong, S. and Ha, T. Real-Time Observation of RecA Filament Dynamics with Single Monomer Resolution. *Cell*, 126(3):515–527, (2006).

184. Tracy, R. B. and Kowalczykowski, S. C. In vitro Selection of Preferred DNA Pairing Sequences by the Escherichia coli RecA Protein. *Genes & Development*, 10(15):1890–1903, (1996).

185. Li, B. S. Sattin, B. D. and Goh, C. Direct and Real-Time Visualization of the Disassembly of a Single RecA-DNA-ATPγS Complex Using AFM Imaging in Fluid. *Nano Letters*, 6(7):1474–1478, (2006).

186. Bork, J. M. Cox, M. M. and Inman, R. B. RecA Protein Filaments Disassemble in the 5' to 3' Direction on Single-Stranded DNA. *The Journal of Biological Chemistry*, 276(49):45740–45743, (2001).

187. Liu, J. Ehmsen, K. T. Heyer, W-D. and Morrical, S. W. Presynaptic Filament Dynamics in Homologous Recombination and DNA Repair. *Critical Reviews in Biochemistry and Molecular Biology*, 46(3):240–270, (2011).

188. De Vlaminck, I. van Loenhout, M. T. J. Zweifel, L. den Blanken, J. Hooning, K. Hage, S. Kerssemakers, J. and Dekker, C. Mechanism of Homology Recognition in DNA Recombination from Dual-Molecule Experiments. *Molecular Cell*, 46(5):616–624, (2012).

189. Mazin, A. V. and Kowalczykowski, S. C. The Function of the Secondary DNA-Binding Site of RecA Protein During DNA Strand Exchange. *EMBO Journal*, 17(4):1161–1168, (1998).

190. Honigberg, S. M. and Radding, C. M. The Mechanics of Winding and Unwinding Helices in Recombination: Torsional Stress Associated with Strand Transfer Promoted by RecA Protein. *Cell*, 54(4):525–532, (1988).

191. Kowalczykowski, S. C. and Krupp, R. A. DNA-Strand Exchange Promoted by RecA Protein in the Absence of ATP: Implications for the Mechanism of Energy Transduction in Protein-Promoted Nucleic Acid Transactions. *Proceedings of the National Academy of Sciences of the United States of America*, 92(8):3478–3482, (1995).

192. Dorfman, K. D. Fulconis, R. Dutreix, M. and Viovy, J. L. Model of RecA-Mediated Homologous Recognition. *Physical Review Letters*, 93(26):3–6, (2004).

193. Kowalczykowski, S. C. Structural biology: Snapshots of DNA repair. *Nature*, 453:4631–466, (2008).

194. Hsieh, P. Camerini-Otero, C. S. and Camerini-Otero, R. D. The Synapsis Event in the Homologous Pairing of DNAs: RecA Recognizes and Pairs Less Than One Helical Repeat of DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 89(14):6492–6496, (1992).

195. Qi, Z. Redding, S. Lee, J. Y. Gibb, B. Kwon, Y. H. Niu, H. Gaines, W. A. Sung, P. and Greene, E. C. DNA Sequence Alignment by Microhomology Sampling During Homologous Recombination. *Cell*, 160(5):856–869, (2015).

196. Radding, C. Synaptic Complex Revisited : A Homologous Recombinase Flips and Switches Bases. *Molecular Cell*, 15(6):846–847, (2004).

197. Zhou, X. and Adzuma, K. DNA Strand Exchange Mediated by the Escherichia coli RecA Protein Initiates in the Minor Groove of Double-Stranded DNA. *Biochemistry*, 36(15):4650–4661, (1997).

198. Adzuma, K. Stable Synapsis of Homologous DNA Molecules Mediated by the Escherichia coli RecA Protein Involves Local Exchange of DNA Strands. *Genes & Development*, 6(9):1679–1694, (1992).

199. Reymer, A. Babik, S. Takahashi, M. Nordén, B. and Beke-Somfai, T. ATP Hydrolysis in the RecADNA Filament Promotes Structural Changes at the ProteinDNA Interface. *Biochemistry*, 54(30):4579-4582, (2015).

200. Menetski, J. P. Bear, D. G. and Kowalczykowski, S. C. Stable DNA Heteroduplex Formation Catalyzed by the Escherichia coli RecA Protein in the Absence of ATP Hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America*, 87(1):21–25, (1990).

201. Rosselli, W. and Stasiak, A. Energetics of RecA-Mediated Recombination Reactions: Without ATP Hydrolysis RecA can Mediate Polar Strand Exchange but is Unable to Recycle. *Journal of Molecular Biology*, 216(2):335–352, (1990).

202. Barzel, A. and Kupiec, M. Finding a Match: How do Homologous Sequences Get Together for Recombination? *Nature Reviews. Genetics*, 9(1):27–37, (2008).

203. Neale, M. J. and Keeney, S. Clarifying the Mechanics of DNA Strand Exchange in Meiotic Recombination. *Nature*, 442(7099):153–158, (2006).

204. Xiao, J. Lee, A. M. and Singleton, S. F. Direct Evaluation of a Kinetic Model for RecA-Mediated DNA-Strand Exchange: the Importance of Nucleic Acid Dynamics and Entropy During Homologous Genetic Recombination. *Chembiochem*, 7(8):1265–1278, (2006).

205. Braun, E. Eichen, Y. Sivan, U. and Ben-Yoseph, G. DNA-Templated Assembly and Electrode Attachment of a Conducting Silver Wire. *Nature*, 391(6669):775–778, (1998).

206. Ellouze, C. Takahashi, M. Wittung, P. Mortensen, K. Schnarr, M. and Norden, B. Evidence for Elongation of the Helical Pitch of the RecA Filament Upon ATP and ADP Binding Using Small-Angle Neutron Scattering. *European Journal of Biochemistry*, 233(2):579–583, (1995).

207. Wittung-stafshede, P. Ellouze, C. Kim, H-K. Mortensen, K. and Takahashi, M. Base Orientation of Second DNA in RecA DNA Filaments Analysis by a Combination of Linear Dichroism and Small Angle Neutron Scattering. *Journal of Biological Chemistry*, 273(25):15682–15686, (1998).

208. Timmins, P. A. Ruigrok, R. W. and DiCapua, E. The Solution Structure of RecA Filaments by Small Angle Neutron Scattering. *Biochimie*, 73(2-3):227–230, (1991).

209. Umemura, K. Komatsu, J. Uchihashi, T. Choi, N. Ikawa, S. Nishinaka, T. Shibata, T. Nakayama, Y. Katsura, S. Mizuno, A. Tokumoto, H. Ishikawa, M. and Kuroda, R. Atomic Force Microscopy of RecA–DNA Complexes Using a Carbon Nanotube Tip. *Biochemical and Biophysical Research Communications*, 281(2):390–395, (2001).

210. Binnig, G. Quate, C. F. and Gerber, C.H. Atomic Force Microscope. *Physical Review Letters*, 56(9):930–933, (1986).

211. Lee, A. J. Interrogation of the RecA Mediated Patterning Mechanism; a Vision of Bioelectronics. *Masters Interim Project Report*, University of Leeds, (2012).

212. McIlwraith, M. J. and West, S. C. The Efficiency of Strand Invasion by Escherichia coli RecA is Dependent Upon the Length and Polarity of ssDNA Tails. *Journal of Molecular Biology*, 305(1):23–31, (2001).

213. Sattin, B. D. and Goh, M. C. Direct Observation of the Assembly of RecA / DNA Complexes by Atomic Force Microscopy. *Biophysical Journal*, 87(5):3430–3436, (2004).

214. Hegner, M. Smith, S. B. and Bustamante, C. Polymerization and Mechanical Properties of Single RecA-DNA Filaments. *Proceedings of the National Academy of Sciences of the United States of America*, 96(18):10109–10114, (1999).

215. Bell, J. C. Plank, J. L. Dombrowski, C. C. and Kowalczykowski, S. C. Direct Imaging of RecA Nucleation and Growth on Single Molecules of SSB-Coated ssDNA. *Nature*, 491:274-278, (2012).

216. Conover, A. J. Danilowicz, C. Gunaratne, R. Coljee, V. W. Kleckner, N. and Prentiss, M. Changes in the Tension in dsDNA Alter the Conformation of RecA Bound to dsDNA-RecA Filaments. *Nucleic Acids Research*, 39(20):8833–8843, (2011).

217. Arata, H. Dupont, A. Miné-Hattab, J. Disseau, L. Renodon-Cornière, A. Takahashi, M. Viovy, J-L. and Cappello, G. Direct Observation of Twisting Steps During Rad51 Polymerization on DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 106(46):19239–19244, (2009).

218. Guo, C. Song, Y. Wang, L. Sun, L. Sun, Y. Peng, C. Liu, Z. Yang, T. and Li, Z. Atomic Force Microscopic Study of Low Temperature Induced Disassembly of RecA-dsDNA Filaments. *The Journal of Physical Chemistry. B*, 112(3):1022–1027, (2008).

219. Guo, C. Li, G. Liu, Z. Sun, L. Sun, Y. Xu, F. Zhang, Y. Yang, T. and Li, Z. Influence of Polyelectrolyte on DNA-RecA Nucleoprotein Filaments: Poly-L-Lysine Used as a Model. *Chemphyschem: a European Journal of Chemical Physics and Physical Chemistry*, 10(9-10):1624–1629, (2009).

220. Lee, A. J. Towards High-Resolution Visualisation of the RecA-Protein Mediated Patterning of DNA-Based Nanostructures. *Masters Dissertation*, University of Leeds, (2012).

221. Klapstein, K. Chou, T. and Bruinsma, R. Physics of RecA-Mediated Homologous Recognition. *Biophysical Journal*, 87(3):1466–1477, (2004).

222. Lee, A. M. Xiao, J. and Singleton, S. F. Origins of Sequence Selectivity in Homologous Genetic Recombination: Insights From Rapid Kinetic Probing of RecA-Mediated DNA Strand Exchange. *Journal of Molecular Biology*, 360(2):343–359, (2006).

223. Nandakumar, D. and Patel, S. S. Finding the Right Match Fast. *Cell*, 160(5):809–811, (2015).

224. Bertucat, G. Lavery, R. and Prévost, C. A Molecular Model for RecA-Promoted Strand Exchange Via Parallel Triple-Stranded Helices. *Biophysical Journal*, 77(3):1562–1576, (1999).

225. Kasas, S. Thomson, N. H. Smith, B. L. Hansma, H. G. Zhu, X. Guthold, M. Bustamante, C. Kool, E. T. Kashlev, M. and Hansma, P. K. Accelerated Publications Escherichia coli RNA Polymerase Activity Observed Using Atomic Force Microscopy. *Biochemistry*, 36(3):461-467, (1997).

226. Guthold, M. Bezanilla, M. Erie, D. A. Jenkins, B. Hansma, H. G. and Bustamante, C. Following the Assembly of RNA Polymerase-DNA Complexes in Aqueous Solutions with the Scanning Force Microscope. *Proceedings of the National Academy of Sciences of the United States of America*, 91(26):12927–12931, (1994).

227. Erie, D. A. Yang, G. L. Schultz, H. C. and Bustamante, C. DNA Bending by Cro Protein in Specific and Nonspecific Complexes: Implications for Protein Site Recognition and Specificity. *Science*, 266(5190):1562–1566, (1994).

228. van Noort, J. Orsini, F. Eker, A. Wyman, C. de Grooth, B. and Greve, J. DNA Bending by Photolyase in Specific and Non-Specific Complexes Studied by Atomic Force Microscopy. *Nucleic Acids Research*, 27(19):3875–3880, (1999).

229. Leuba, S. H. Yang, G. Robert, C. Samori, B. van Holde, K. Zlatanova, J. and Bustamante, C. Three-Dimensional Structure of Extended Chromatin Fibers as Revealed by Tapping-Mode Scanning Force Microscopy. *Proceedings of the National Academy of Sciences of the United States of America*, 91(24):11621–11625, (1994).

230. Sorel, I. Piétrement, O. Hamon, L. Baconnais, S. Le Cam, E. and Pastré, D. The EcoRI-DNA Complex as a Model for Investigating Protein-DNA Interactions by Atomic Force Microscopy. *Biochemistry*, 45(49):14675–14682, (2006).

231. Kodera, N. Yamamoto, D. Ishikawa, R. and Ando, T. Video Imaging of Walking Myosin V by High-Speed Atomic Force Microscopy. *Nature*, 468(7320):72–76, (2010).

232. Shibata, M. Yamashita, H. Uchihashi, T. Kandori, H. and Ando, T. High-Speed Atomic Force Microscopy Shows Dynamic Molecular Processes in Photoactivated Bacteriorhodopsin. *Nature Nanotechnology*, 5(3):208–212, (2010).

233. Uchihashi, T. Iino, R. Ando, T. and Noji, H. High-Speed Atomic Force Microscopy Reveals Rotary Catalysis of Rotorless F1-ATPase. *Science*, 333(6043):755–759, (2011).

234. Rajendran, A. Endo, M. Hidaka, K. and Sugiyama, H. Direct and Single-Molecule Visualization of the Solution-State Structures of G-Hairpin and G-Triplex Intermediates. *Angewandte Chemie*, 126(16):4191–4196, (2014).

235. Rajendran, A. Endo, M. Hidaka, K. Tran, P. L. T. Mergny, J-L. and Sugiyama, H. Controlling the Stoichiometry and Strand Polarity of a Tetramolecular G-Quadruplex Structure by Using a DNA Origami Frame. *Nucleic Acids Research*, 41(18):8738–8747, (2013).

236. Suzuki, Y. Endo, M. Katsuda, Y. Ou, K. Hidaka, K. and Sugiyama, H. DNA Origami Based Visualization System for Studying Site-Specific Recombination Events. *Journal of the American Chemical Society*, 136(1):211–218, (2014).

237. Bath, J. and Turberfield, A. J. DNA Nanomachines. *Nature Nanotechnology*, 2(5):275–284, (2007).

238. Stojanovic, M. N. Stefanovic, D. and Rudchenko, S. Exercises in Molecular Computing. *Accounts of Chemical Research*, 47(6):1845–1852, (2014).

239. Lee, A. J. and Walti, C. Studying Biologically Templated Materials with Atomic Force Microscopy. In *Nanomaterials: A Guide to Fabrication and Applications*, edited by Krishnamoorthy, S. CRC Press, (2015).

240. Santos, N. C. and Castanho, M. A. R. B. An Overview of the Biophysical Applications of Atomic Force Microscopy. *Biophysical Chemistry*, 107(2):133–149, (2004).

241. García, R. *Amplitude Modulation Atomic Force Microscopy*. John Wiley & Sons, (2011).

242. Martin, Y. Williams, C. C. and Wickramasinghe, H. K. Atomic Force Microscope Force Mapping and Profiling on a Sub 100-AŁ Scale. *Journal of Applied Physics*, 61(10):4723, (1987).

243. Zhong, Q. Inniss, D. Kjoller, K. and Elings, V.B. Fractured Polymer / Silica Fiber Surface Studied by Tapping Mode Atomic Force Microscopy. *Surface Science Letters*, 290:L688–L692, (1993).

244. Hansma, H. G. Browne, K. A. Bezanilla, M. and Bruiceg, T. C. Bending and Straightening of DNA Induced by the Same Ligand: Characterization with the Atomic Force Microscope. *Biochemistry*, 33(28):8436–8441, (1994).

245. Schaffer, T. E. Cleveland, J. P. Ohnesorge, F. Walters, D. A. and Hansma, P. K. Studies of Vibrating Atomic Force Microscope Cantilevers in Liquid. *Journal of Applied Physics*, 80(7):3622, (1996).

246. Tortonese, M. Barrett, R. C. and Quate, C. F. Atomic Resolution with an Atomic Force Microscope Using Piezoresistive Detection. *Applied Physics Letters*, 62(8):834, (1993).

247. Meyer, G. and Amer, N. M. Novel Optical Approach to Atomic Force Microscopy. *Applied Physics Letters*, 53(12):1045, (1988).

248. Argento, C. and French, R. H. Parametric Tip Model and ForceDistance Relation for Hamaker Constant Determination from Atomic Force Microscopy. *Journal of Applied Physics*, 80(11):6081, (1996).

249. Hertz, H. Über die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik*, 92:156–171, (1881).

250. Derjaguin, B. V. Muller, V. M. and Toporov, Y. U. P. Effect of Contact Deformations on the Adhesion of Particles. *Progress in Surface Science*, 45(1-4):131–143, (1994).

251. Johnson, K. L. and Greenwood, J. A. An Adhesion Map for the Contact of Elastic Spheres. *Journal of Colloid and Interface Science*, 333(192):326–333, (1997).

252. Johnson, K. L. Kendall, K. and Roberts, A. D. Surface Energy and the Contact of Elastic Solids. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical*, 324(1558):301–313, (1971).

253. Fisher, R. and Israelachvili, J. Experimental Studies on the Applicability of the Kelvin Equation to Highly Curved Concave Menisci. *Journal of Colloid and Interface Science*, 80(2):528–541, (1981).

254. Israelachvili, J. *Intermolecular and Surface Forces*. Elsevier, London, 3rd edition, (2011).

255. Butt, H-J. Cappella, B. and Kappl, M. Force Measurements with the Atomic Force Microscope: Technique, Interpretation and Applications. *Surface Science Reports*, 59(1-6):1–152, (2005).

256. Müller, D. J. Fotiadis, D. Scheuring, S. Müller, S. A. and Engel, A. Electrostatically Balanced Subnanometer Imaging of Biological Specimens by Atomic Force Microscope. *Biophysical Journal*, 76(2):1101–1111, (1999).

257. Sotres, J. and Baró, A. M. AFM Imaging and Analysis of Electrostatic Double Layer Forces on Single DNA Molecules. *Biophysical Journal*, 98(9):1995–2004, (2010).

258. Morris, V. J. Kirby, A. R. and Gunning, A. P. *Atomic Force Microscopy for Biologists*. Imperial College Press, London, 2nd Edition, (2010).

259. Fukuma, T. and Jarvis, S. P. Development of Liquid-Environment Frequency Modulation Atomic Force Microscope with Low Noise Deflection Sensor for Cantilevers of Various Dimensions. *Review of Scientific Instruments*, 77(4):043701, (2006).

260. Kilpatrick, J. I. Loh, S.-H. and Jarvis, S. P. Directly Probing the Effects of Ions on Hydration Forces at Interfaces. *Journal of the American Chemical Society*, 135(7):2628–2634, (2013).

261. Putman, C. A. J. Van der Werf, K. O. De Grooth, B. G. Van Hulst, N. F. and Greve, J. Tapping Mode Atomic Force Microscopy in Liquid. *Applied Physics Letters*, 64(18):2454, (1994).

262. Ando, T. High-Speed Atomic Force Microscopy Coming of Age. *Nanotechnology*, 23(6):062001, (2012).

263. Basak, S. Raman, A. and Garimella, S. V. Hydrodynamic Loading of Microcantilevers Vibrating in Viscous Fluids. *Journal of Applied Physics*, 99(11):114906, (2006).

264. Basak, S. and Raman, A. Dynamics of Tapping Mode Atomic Force Microscopy in Liquids: Theory and Experiments. *Applied Physics Letters*, 91(6):064107, (2007).

265. Green, C. P. and Sader, J. E. Frequency Response of Cantilever Beams Immersed in Viscous Fluids Near a Solid Surface with Applications to the Atomic Force Microscope. *Journal of Applied Physics*, 98(11):114913, (2005).

266. Raman, A. Melcher, J. and Tung, R. Cantilever Dynamics in Atomic Force Microscopy. *Nano Today*, 3(1):20–27, (2008).

267. Giessibl, F. J. Advances in Atomic Force Microscopy. *Reviews of Modern Physics*, 75(3):949–983, (2003).

268. Fukuma, T. Kobayashi, K. Matsushige, K. and Yamada, H. True Molecular Resolution in Liquid by Frequency-Modulation Atomic Force Microscopy. *Applied Physics Letters*, 86(19):193108, (2005).

269. Fukuma, T. Wideband Low-Noise Optical Beam Deflection Sensor with Photothermal Excitation for Liquid-Environment Atomic Force Microscopy. *The Review of Scientific Instruments*, 80(2):023707, (2009).

270. Ando, T. Uchihashi, T. and Fukuma, T. High-Speed Atomic Force Microscopy for Nano-Visualization of Dynamic Biomolecular Processes. *Progress in Surface Science*, 83(7-9):337–437, (2008).

271. Martínez, L. Tello, M. Díaz, M. Román, E. Garcia, R. and Huttel, Y. Aspect-Ratio and Lateral-Resolution Enhancement in Force Microscopy by Attaching Nanoclusters Generated by an Ion Cluster Source at the End of a Silicon Tip. *The Review of Scientific Instruments*, 82(2):023710, (2011).

272. Stevens, R. M. New Carbon Nanotube AFM Probe Technology. *Materials Today*, 12(10):42–45, (2009).

273. Möller, C. Allen, M. Elings, V. Engel, A. and Müller, D. J. Tapping-Mode Atomic Force Microscopy Produces Faithful High-Resolution Images of Protein Surfaces. *Biophysical Journal*, 77(2):1150–1158, (1999).

274. Mullin, N. and Hobbs, J. K. Torsional Resonance Atomic Force Microscopy in Water. *Applied Physics Letters*, 92(5):053103, (2008).

275. Mullin, N. Vasilev, C. Tucker, J. D. Hunter, C. N. Weber, C. H. M. and Hobbs, J. K. Torsional Tapping Atomic Force Microscopy Using T-Shaped Cantilevers. *Applied Physics Letters*, 94(17):173109, (2009).

276. Mullin, N. and Hobbs, J. Direct Imaging of Polyethylene Films at Single-Chain Resolution with Torsional Tapping Atomic Force Microscopy. *Physical Review Letters*, 107(19):1–5, (2011).

277. Kaur, P. Qiang, F. Fuhrmann, A. Ros, R. Kutner, L. O. Schneeweis, L. A. Navoa, R. Steger, K. Xie, L. Yonan, C. Abraham, R. Grace, M. J. and Lindsay, S. Antibody-Unfolding and Metastable-State Binding in Force Spectroscopy and Recognition Imaging. *Biophysical Journal*, 100(1):243–250, (2011).

278. Gunning, A. P. Chambers, S. Pin, C. Man, A. L. Morris, V. J. and Nicoletti, C. Mapping Specific Adhesive Interactions on Living Human Intestinal Epithelial Cells With Atomic Force Microscopy. *FASEB Journal*, 22(7):2331–2339, (2008).

279. Hoffmann, T. and Dougan, L. Single Molecule Force Spectroscopy Using Polyproteins. *Chemical Society Reviews*, 41(14):4781–4796, (2012).

280. Cappella, B. and Dietler, G. Force-Distance Curves by Atomic Force Microscopy. *Surface Science Reports*, 34(1-3):1–104, (1999).

281. Morsi, S. M. Pakzad, A. Amin, A. Yassar, R. S. and Heiden, P. A. Chemical and Nanomechanical Analysis of Rice Husk Modified by ATRP-grafted Oligomer. *Journal of Colloid and Interface Science*, 360(2):377–385, (2011).

282. Kaemmer, S.B. Application Note # 133 Introduction to Brukers ScanAsyst and PeakForce Tapping AFM Technology. Technical Report, Bruker Nano Surfaces Division, Santa Barbara, (2011).

283. Alsteens, D. Dupres, V. Yunus, S. Latgé, J.-P. Heinisch, J. J. and Dufrêne, Y. F. High-Resolution Imaging of Chemical and Biological Sites on Living Cells Using Peak Force Tapping Atomic Force Microscopy. *Langmuir*, 28(49):16738–16744, (2012).

284. Trtik, P. Kaufmann, J. and Volz, U. On the Use of Peak-Force Tapping Atomic Force Microscopy for Quantification of the Local Elastic Modulus in Hardened Cement Paste. *Cement and Concrete Research*, 42(1):215–221, (2012).

285. Pittenger, B. Slade, A. Berquand, A. Milani, P. Boudaoud, A. Hamant, O. and Radmacher, M. Application Note #141: Toward Quantitative Nanomechanical Measurements on Live Cells with PeakForce QNM. Technical Report, Bruker Nano Surfaces Division, Santa Barbara (2013).

286. Hansma, H. G. Surface Biology of DNA by Atomic Force Microscopy. *Annual Review of Physical Chemistry*, 52:71–92, (2001).

287. Florin, E. L. Moy, V. T. and Gaub, H. E. Adhesion Forces Between Individual Ligand-Receptor Pairs. *Science*, 264(5157):415–417, (1994).

288. Moy, V. T. Florin, E. L. and Gaub, H. E. Intermolecular Forces and Energies Between Ligands and Receptors. *Science*, 266(5183):257–259, (1994).

289. Ando, T. Kodera, N. Takai, E. Maruyama, D. Saito, K. and Toda, A. A High-Speed Atomic Force Microscope for Studying Biological Macromolecules. *Proceedings of the National Academy of Sciences of the United States of America*, 98(22):12468–12472, (2001).

290. Yamamoto, D. Uchihashi, T. Kodera, N. Yamashita, H. Nishikori, S. Ogura, T. Shibata, M. and Ando, T. High-Speed Atomic Force Microscopy Techniques for Observing Dynamic Biomolecular Processes. *Methods in Enzymology*, 475(10):541–564, (2010).

291. Ando, T. Uchihashi, T. and Kodera, N. High-Speed Atomic Force Microscopy. *Japanese Journal of Applied Physics*, 51(8):08KA02, (2012).

292. Ando, T. Kodera, N. Naito, Y. Kinoshita, T. Furuta, K. and Toyoshima, Y. Y. A High-Speed Atomic Force Microscope for Studying Biological Macromolecules in Action. *Chemphyschem*, 4(11):1196–1202, (2003).

293. Ando, T. Kodera, N. Uchihashi, T. Miyagi, A. Nakakita, R. Yamashita, H. and Matada, K. High-Speed Atomic Force Microscopy for Capturing Dynamic Behavior of Protein Molecules at Work. *e-Journal of Surface Science and Nanotechnology*, 3:384–392, (2005).

294. Uchihashi, T. Kodera, N. and Ando, T. Guide to Video Recording of Structure Dynamics and Dynamic Processes of Proteins by High-Speed Atomic Force Microscopy. *Nature Protocols*, 7(6):1193–1206, (2012).

295. Manalis, S. R. Minne, S. C. Atalar, A. and Quate, C. F. Interdigital Cantilevers for Atomic Force Microscopy. *Applied Physics Letters*, 69(25):3944, (1996).

296. Sulchek, T. Minne, S. C. Adams, J. D. Fletcher, D. A. Atalar, A. Quate, C. F. and Adderton, D. M. Dual Integrated Actuators for Extended Range High Speed Atomic Force Microscopy. *Applied Physics Letters*, 75(11):1637, (1999).

297. Sulchek, T. Yaralioglu, G. G. Quate, C. F. and Minne, S. C. Characterization and Optimization of Scan Speed for Tapping-Mode Atomic Force Microscopy. *Review of Scientific Instruments*, 73(8):2928, (2002).

298. Adams, J.D. Manning, L. Rogers, B. Jones, M. and Minne, S.C. Self-Sensing Tapping Mode Atomic Force Microscopy. *Sensors and Actuators A: Physical*, 121(1):262–266, (2005).

299. Rogers, B. Manning, L. Sulchek, T. and Adams, J. D. Improving Tapping Mode Atomic Force Microscopy with Piezoelectric Cantilevers. *Ultramicroscopy*, 100(3-4):267–276, (2004).

300. Rogers, B. Sulchek, Murray, T. K. York, D. Jones, M. Manning, L. Malekos, S. Beneschott, B. Adams, J. D. Cavazos, H. and Minne, S. C. High Speed Tapping Mode Atomic Force

Microscopy in Liquid Using an Insulated Piezoelectric Cantilever. *Review of Scientific Instruments*, 74(11):4683, (2003).

301. Humphris, A. D. L. Miles, M. J. and Hobbs, J. K. A Mechanical Microscope: High-Speed Atomic Force Microscopy. *Applied Physics Letters*, 86(3):034106, (2005).

302. Hobbs, J. K. Vasilev, C. and Humphris, A. D. L. Real Time Observation of Crystallization in Polyethylene Oxide with Video Rate Atomic Force Microscopy. *Polymer*, 46(23):10226–10236, (2005).

303. Walters, D. A. Cleveland, J. P. Thomson, N. H. Hansma, P. K. Wendman, M. A. Gurley, G. and Elings, V. Short Cantilevers for Atomic Force Microscopy. *Review of Scientific Instruments*, 67(10):3583, (1996).

304. Viani, M. B. Schaffer, T. E. Chand, A. Rief, M. Gaub, H. E. and Hansma, P. K. Small Cantilevers for Force Spectroscopy of Single Molecules. *Journal of Applied Physics*, 86(4):2258, (1999).

305. Viani, M. B. Schaffer, T. E. Paloczi, G. T. Pietrasanta, L. I. Smith, B. L. Thompson, J. B. Richter, M. Rief, M. Gaub, H. E. Plaxco, K. W. Cleland, A. N. Hansma, H. G. and Hansma, P. K. Fast Imaging and Fast Force Spectroscopy of Single Biopolymers with a New Atomic Force Microscope Designed for Small Cantilevers. *Review of Scientific Instruments*, 70(11):4300, (1999).

306. Fukuma, T. Okazaki, Y. Kodera, N. Uchihashi, T. and Ando, T. High Resonance Frequency Force Microscope Scanner Using Inertia Balance Support. *Applied Physics Letters*, 92(24):243119, (2008).

307. Kodera, N. Yamashita, H. and Ando, T. Active Damping of the Scanner for High-Speed Atomic Force Microscopy. *Review of Scientific Instruments*, 76(5):053708, (2005).

308. Kodera, N. Sakashita, M. and Ando, T. Dynamic Proportional-Integral-Differential Controller for High-Speed Atomic Force Microscopy. *Review of Scientific Instruments*, 77(8):083704, (2006).

309. Fantner, G. E. Hegarty, P. Kindt, J. H. Schitter, G. Cidade, G. A. G. and Hansma, P. K. Data Acquisition System for High Speed Atomic Force Microscopy. *Review of Scientific Instruments*, 76(2):026118, (2005).

310. Fantner, G. E. Schitter, G. Kindt, J. H. Ivanov, T. Ivanova, K. Patel, R. Holten-Andersen, N. Adams, J. Thurner, P. J. Rangelow, I. W. and Hansma, P. K. Components for High Speed Atomic Force Microscopy. *Ultramicroscopy*, 106(8-9):881–887, (2006).

311. Hansma, P. K. Schitter, G. Fantner, G. E. and Prater, C. High-speed atomic force microscopy. *Science*, 314(5799):601–602, (2006).

312. Sharma, R. *Protein-Mediated Patterning of DNA Scaffolds for Nanoscale Electronics*. PhD thesis, University of Leeds, (2009).

313. Sambrook, J. and Russell, D. W. *Molecular Cloning: A Laboratory Manual. Vol 1 - 3.* Cold Spring Harbor Laboratory Press, (2001).

314. Douglas, S. M. Dietz, H. Liedl, T. Högberg, B. Graf, F. and Shih, W. M. Self-Assembly of DNA into Nanoscale Three-Dimensional Shapes. *Nature*, 459(7245):414–418, (2009).

315. Sharma, R. Davies, A. G. and Wälti, C. Nanoscale Programmable Sequence-Specific Patterning of DNA Scaffolds Using RecA Protein. *Nanotechnology*, 23(36):365301, (2012).

316. Paulus. B. F. and Bryant, F. R. Time-Dependent Inhibition of RecA Protein-Catalyzed ATP Hydrolysis by ATPγS: Evidence for a Rate-Determining Isomerization of the RecA-ssDNA Complex. *Biochemistry*, 36(25):7832–7838, (1997).

317. Rice, M. C. Bruner, M. Czymmek, K. and Kmiec, E. B. In vitro and in vivo Nucleotide Exchange Directed by Chimeric RNA/DNA Oligonucleotides in Saccharomyces cerevisae. *Molecular Microbiology*, 40(4):857–868, (2001).

318. Weiner, A. Zauberman, N. and Minsky, A. Recombinational DNA Repair in a Cellular Context: a Search for the Homology Search. *Nature Reviews. Microbiology*, 7(10):748–755, (2009).

319. Billingsley, D. J. Lee, A. J. Johansson, N. A. B. Walton, A. Stanger, L. Crampton, N. Bonass, W. A. and Thomson, N. H. Patchiness of Ion-Exchanged Mica Revealed by DNA Binding Dynamics at Short Length Scales. *Nanotechnology*, 25(2):025704, (2014).

320. Lee, A. J. Szymonik, M. Hobbs, J. K. and Wälti, C. Tuning the Translational Freedom of DNA for High Speed AFM. *Nano Research*, 8(6):1811-1821, (2015).

321. Yancey-Wrona, J. E. and Camerini-Otero, R. D. The Search for DNA Homology Does Not Limit Stable Homologous Pairing Promoted by RecA Protein. *Current Biology*, 5(10):1149–1158, (1995).

322. Shi, W-X. and Larson, R. G. Atomic Force Microscopic Study of Aggregation of RecA-DNA Nucleoprotein Filaments into Left-Handed Supercoiled Bundles. *Nano Letters*, 5(12):2476–2481, (2005).

323. Lesterlin, C. Ball, G. Schermelleh, L. and Sherratt, D. J. RecA Bundles Mediate Homology Pairing Between Distant Sisters During DNA Break Repair. *Nature*, 506(7487):249–253, (2014).

324. Balmer, T. E. Christenson, H. K. Spencer, N. D. and Heuberger, M. The Effect of Surface Ions on Water Adsorption to Mica. *Langmuir*, 24(4):1566–1569, (2008).

325. Hansma, H. G. and Laney, D. E. DNA Binding to Mica Correlates with Cationic Radius: Assay by Atomic Force Microscopy. *Biophysical Journal*, 70(4):1933–1939, (1996).

326. Lyubchenko, Y. L. and Shlyakhtenko, L. S. AFM for Analysis of Structure and Dynamics of DNA and Protein-DNA Complexes. *Methods*, 47(3):206–213, (2009).

327. Josephs, E. A. and Ye, T. Electric-Field Dependent Conformations of Single DNA Molecules on a Model Biosensor Surface. *Nano Letters*, 12(10):5255–5261, (2012).

328. Josephs, E. A. and Ye, T. A Single-Molecule View of Conformational Switching of DNA Tethered to a Gold Electrode. *Journal of the American Chemical Society*, 134(24):10021–10030, (2012).

329. Thomson, N. H. Fritz, M. Radmacher, M. Cleveland, J. P. Schmidt, C. F. and Hansma, P. K. Protein Tracking and Detection of Protein Motion Using Atomic Force Microscopy. *Biophysical Journal*, 70(5):2421–2431, (1996).

330. Pietrement, O. Pastre, D. Fusil, S. Jeusset, J. David, M-O. Landousy, F. Hamon, L. Zozime, A. and Le Cam, E. Reversible Binding of DNA on $NiCl_2$-Treated Mica by Varying the Ionic Strength. *Langmuir*, 19(15):2536–2539, (2003).

331. Abel, G. R. Josephs, E. A. Luong, N. and Ye, T. A Switchable Surface Enables Visualization of Single DNA Hybridization Events with Atomic Force Microscopy. *Journal of the American Chemical Society*, 135:6399–6402, (2013).

332. Erdmann, M. David, R. Fornof, A. and Gaub, H. E. Electrically Controlled DNA Adhesion. *Nature Nanotechnology*, 5(2):154–159, (2010).

333. Hsueh, C. Chen, H. Gimzewski, J. K. Reed, J. and Abdel, T. M. Localized Nanosopic Surface Measurements of Nickel-Modified Mica for Single Molecule DNA Sequence Sampling. *ACS Applied Material Interfaces*, 2(11):3249–3256, (2011).

334. Zheng, J. Li, Z. Wu, A. and Zhou, H. AFM Studies of DNA Structures on Mica in the Presence of Alkaline Earth Metal Ions. *Biophysical Chemistry*, 104(1):37–43, (2003).

335. Pastré, D. Piétrement, O. Fusil, S. Landousy, F. Jeusset, J. David, M-O. Hamon, L. Le Cam, E. and Zozime, A. Adsorption of DNA to Mica Mediated by Divalent Counterions: a Theoretical and Experimental Study. *Biophysical Journal*, 85(4):2507–2518, (2003).

336. Ellis, J. S. Abdelhady, H. G. Allen, S. Davies, M. C. Roberts, C. J. Tendler, S. J. B. and Williams, P. M. Direct Atomic Force Microscopy Observations of Monovalent Ion Induced Binding of DNA to Mica. *Journal of Microscopy*, 215(Pt 3):297–301, (2004).

337. Pastré, D. Hamon, L. Landousy, F. Sorel, I. David, M-O. Zozime, A. Le Cam, E. and Piétrement, O. Anionic Polyelectrolyte Adsorption on Mica Mediated by Multivalent Cations: a Solution to DNA Imaging by Atomic Force Microscopy Under High Ionic Strengths. *Langmuir*, 22(15):6651–6660, (2006).

338. Jahnen-Dechent, W. and Ketteler, M. Magnesium Basics. *Clinical Kidney Journal*, 5:i3–i14, (2012).

339. Smith, M. B. Li, H. Shen, T. Huang, X. Yusuf, E. and Vavylonis, D. Segmentation and Tracking of Cytoskeletal Filaments Using Open Active Contours. *Cytoskeleton*, 67(11):693–705, (2010).

340. Lyubchenko, Y. L. Preparation of DNA and Nucleoprotein Samples for AFM Imaging. *Micron*, 42(2):196–206, (2011).

341. Guthold, M. Zhu, X. Rivetti, C. Yang, G. Thomson, N. H. Kasas, S. Hansma, H. G. Smith, B. Hansma, P. K. and Bustamante, C. Direct Observation of One-Dimensional Diffusion and Transcription by Escherichia coli RNA Polymerase. *Biophysical Journal*, 77(4):2284–2294, (1999).

342. Bennink, M. L. Nikova, D. N. van der Werf, K. O. and Greve, J. Dynamic Imaging of Single DNAProtein Interactions Using Atomic Force Microscopy. *Analytica Chimica Acta*, 479(1):3–15, (2003).

343. Suzuki, Y. Higuchi, Y. Hizume, K. Yokokawa, M. Yoshimura, S. H. Yoshikawa, K. and Takeyasu, K. Molecular Dynamics of DNA and Nucleosomes in Solution Studied by Fast-Scanning Atomic Force Microscopy. *Ultramicroscopy*, 110(6):682–688, (2010).

344. Kobayashi, M. Sumitomo, K. and Torimitsu, K. Real-Time Imaging of DNA-Streptavidin Complex Formation in Solution Using a High-Speed Atomic Force Microscope. *Ultramicroscopy*, 107(2-3):184–190, (2007).

345. Rajendran, A. Endo, M. Hidaka, K. and Sugiyama, H. Direct and Real-Time Observation of Rotary Movement of a DNA Nanomechanical Device. *Journal of the American Chemical Society*, 135(3):1117–1123, (2013).

346. Endo, M. Katsuda, Y. Hidaka, K. and Sugiyama, H. Regulation of DNA Methylation Using Different Tensions of Double Strands Constructed in a Defined DNA Nanostructure. *Journal of the American Chemical Society*, 132(5):1592–1597, (2010).

347. Nimse, S. Song, K. Sonawane, M. Sayyed, D. and Kim, T. Immobilization Techniques for Microarray: Challenges and Applications. *Sensors*, 14(12):22208–22229, (2014).

348. Fischer, F. Henning-Knechtel, A. and Mertig, M. Investigating the Aggregation Behaviour of DNA Origami Frames. *Physica Status Solidi (a)*, 212(6):1375–1381, (2015).

349. Slutsky, M. and Mirny, L. A. Kinetics of Protein-DNA Interaction: Facilitated Target Location in Sequence-Dependent Potential. *Biophysical Journal*, 87(6):4021–4035, (2004).

350. Kates-Harbeck, J. Tilloy, A. and Prentiss, M. Simplified Biased Random Walk Model for RecA-Protein-Mediated Homology Recognition Offers Rapid and Accurate Self-Assembly of Long Linear Arrays of Binding Sites. *Physical Review E*, 88(1):012702, (2013).

351. Jiang, L. and Prentiss, M. RecA-Mediated Sequence Homology Recognition as an Example of How Searching Speed in Self-Assembly Systems can be Optimized by Balancing Entropic and Enthalpic Barriers. *Physical Review E*, 90(2):022704, (2014).

352. Walz, T. Tittmann, P. Fuchs, K. H. Müller, D. J. Smith, B. L. Agre, P. Gross, H. and Engel, A. Surface Topographies at Subnanometer-Resolution Reveal Asymmetry and Sidedness of Aquaporin-1. *Journal of Molecular Biology*, 264(5):907–918, (1996).

353. Müller, D. J. Sass, H. J. Müller, S. A. Büldt, G. and Engel, A. Surface Structures of Native Bacteriorhodopsin Depend on the Molecular Packing Arrangement in the Membrane. *Journal of Molecular Biology*, 285(5):1903–1909, (1999).

354. Müller, D. J. Janovjak, H. Lehto, T. Kuerschner, L. and Anderson, K. Observing Structure, Function and Assembly of Single Proteins by AFM. *Progress in Biophysics and Molecular Biology*, 79(1-3):1–43, (2002).

355. Engel, A. and Gaub, H. E. Structure and Mechanics of Membrane Proteins. *Annual Review of Biochemistry*, 77:127–148, (2008).

356. Mou, J. Czajkowsky, D. M. Zhang, Y. and Shao, Z. High-Resolution Atomic Force Microscopy of DNA: the Pitch of the Double Helix. *FEBS Letters*, 371(3):279-282, (1995).

357. Leung, C. Bestembayeva, A. Thorogate, R. Stinson, J. Pyne, A. Marcovich, C. Yang, J. Drechsler, U. Despont, M. Jankowski, T. Tschöpe, M. and Hoogenboom, B. W. Atomic Force Microscopy with Nanoscale Cantilevers Resolves Different Structural Conformations of the DNA Double Helix. *Nano Letters*, 12(7):3846–3850, (2012).

358. Santos, S. Barcons, V. Christenson, H. K. Billingsley, D. J. Bonass, W. A. Font, J. and Thomson, N. H. Stability, Resolution, and Ultra-Low Wear Amplitude Modulation Atomic Force Microscopy of DNA: Small Amplitude Small Set-Point Imaging. *Applied Physics Letters*, 103(6):063702, (2013).

359. Ido, S. Kimura, K. Oyabu, N. Kobayashi, K. Tsukada, M. Matsushige, K. and Yamada, H. Beyond the Helix Pitch: Direct Visualization of Native DNA in Aqueous Solution. *ACS Nano*, 7(2):1817-1822, (2013).

360. Suzuki, K. Kitamura, S-I. Tanaka, S. Kobayashi, K. and Yamada, H. Development of High-Resolution Imaging of SolidLiquid Interface by Frequency Modulation Atomic Force Microscopy. *Japanese Journal of Applied Physics*, 49(8):08LB12, (2010).

361. Pyne, A. Thompson, R. Leung, C. Roy, D. and Hoogenboom, B. W. Single-Molecule Reconstruction of Oligonucleotide Secondary Structure by Atomic Force Microscopy. *Small*, 10(16):3257–3261, (2014).

362. Bai, X-C. Martin, T. G. Scheres, S. H. W. and Dietz, H. Cryo-EM Structure of a 3D DNA-Origami Object. *Proceedings of the National Academy of Sciences of the United States of America*, 109:20012–20017, (2012).

363. Lin, C. Jungmann, R. Leifer, A. M. Li, C. Levner, D. Church, G. M. Shih, W. M. and Yin, P. Submicrometre Geometrically Encoded Fluorescent Barcodes Self-Assembled from DNA. *Nature Chemistry*, 4(10):832–839, (2012).

364. Story, R. M. Weber, I. T. and Steitz, T. A. The Structure of the E. coli RecA Protein Monomer and Polymer. *Nature*, 355(6358):318–325, (1992).

365. Dunn, K. Chrysogelos, S. and Griffith, J. Electron Microscopic Visualization of RecA-DNA Filaments: Evidence for a Cyclic Extension of Duplex DNA. *Cell*, 28(4):757-765, (1982).

366. Travaglini, G. Rohrer, H. and Stoll, E. Scanning Tunneling Microscopy of RecA-DNA Complexes. *Physica Scripta*, 38:309–314, (1988).

367. Amrein, M. Stasiak, A. Gross, H. Stoll, E. and Travaglini, G. Scanning Tunneling Microscopy of RecA-DNA Complexes Coated with a Conducting Film. *Science*, 240(4851):514–516, (1988).

368. Umemura, K. Ikawa, S. Nishinaka, T. Shibata, T. and Kuroda, R. Imaging the RecA-DNA Complex by Atomic Force Microscopy. *Nucleic Acids Symposium Series*, 42:235–236, (1999).

369. Kim, D. N. Kilchherr, F. Dietz, H. and Bathe, M. Quantitative Prediction of 3D Solution Shape and Flexibility of Nucleic Acid Nanostructures. *Nucleic Acids Research*, 40(7):2862–2868, (2012).

370. Endo, M. and Sugiyama, H. Single-Molecule Imaging of Dynamic Motions of Biomolecules in DNA Origami Nanostructures Using High-Speed Atomic Force Microscopy. *Accounts of Chemical Research*, 47(6):1645–1653, (2014).